

分类号 _____

密级 _____

UDC _____

编号 _____

中南大學

CENTRAL SOUTH UNIVERSITY

硕士学位论文

论文题目 基于计算机视觉的体表三维测量

技术研究

学科、专业 生物医学工程

研究生姓名 张必明

指导老师及

专业技术职务 何继善 教授、熊平 教授

DISSERTATION

The research of human body surface 3D measurement technology based on computer vision

Specialty: Biomedical Engineering

Master Degree Candidate: Zhang Biming

Supervisor: Prof. He Jishan Prof. Xiong Ping

School of Info-physics and Engineering

Central South University

ChangSha Hunan P.R.C

April 2006

摘要

体表测量是医学、法医学领域的重要课题之一。医疗诊断中确定烧伤区域面积、瘢痕面积；法医进行体表损伤的临床司法鉴定时，需要测量体表损伤区域的面积、长度、角度等作为鉴定的重要指标。传统方法一般采用直接体表估测或照片辅助测量，存在测量不准确、操作不方便、难以重复测量等缺陷。而基于摄影光学、数字图像处理与计算机视觉的测量方法能够对体表的任意区域进行三维重建，使进一步三维测量体表的面积、长度测量成为可能。

基于计算机视觉的三维测量技术是计算机视觉的重要研究方向，近年来，随着工业化的发展，三维测量技术在工业领域的应用范围越来越广；同时随着医学相关领域技术的不断前进，医学方面的三维测量需求越来越强。计算机视觉的三维测量技术是一种非接触、以软件计算为主的测量技术，具有其独特的优势。

本文从研究计算机视觉、摄影测量技术的发展开始，研究分析三维测量的方法及相关基础理论，逐渐展开深入到测量过程中一些关键技术的算法研究。以测量的过程为线索进行研究，以构建先进、完整的三维测量系统为最终目标。

关键词： 三维测量，计算机视觉，摄像机标定，体表三维重建

ABSTRACT

Human body surface measurement is an important research subject in clinical and forensic medical field. Doctors need to measure area of burns and scars for clinical diagnosis. Forensic doctors need to measure the area, length and angle of wounded and injured area to judge how serious of the damage for forensic judgement. Traditional methods to measure area and length usually directly measure on human body surface or via the photo to aided measure. There are many disadvantages using the traditional metrical methods, such as handle inconveniently, the result not accurate enough, difficult to re-measurement etc. The methods that based on photogrammetry, digital image processing and computer vision could implement 3D reconstruction of the human body surface and make 3D based measurement be possible.

3D measurement technology based on computer vision is an important orientation in computer vision research. Recently, the application of 3D measurement technology widely applied in industrial and medical area. It has some superior features as non-touch, implement measurement via software algorithms mainly means less cost than special devices.

This paper research from the development of the computer vision and measurement technology, then study the associated methods and theory of the 3D measurement, futherly spread and go deep into some key-technology and algorithms in process of measurement. The research is following the clue of the process of 3D measurement. The goal of the research is to build a practical, advanced and integrated 3D measurement system.

Key WORDS 3D measurement, computer vision, camera calibration, human body surface 3D reconstruction

目 录

第一章 绪论	1
1.1 引言	1
1.2 三维测量技术综述	3
1.2.1 三维测量技术发展综述	3
1.2.2 三维测量方法综述	5
1.3 论文的主要内容及组织结构	7
第二章 计算机视觉三维测量理论基础	9
2.1 射影几何	9
2.1.1 射影空间	9
2.1.2 射影变换	11
2.2 摄像机成像模型	12
2.2.1 预备知识	12
2.2.2 线性摄像机模型	14
2.2.3 摄像机畸变	16
2.3 对极几何与基本矩阵	17
2.3.1 对极几何	17
2.3.2 基本矩阵	18
2.4 本章小结	20
第三章 计算机视觉三维测量方法研究	21
3.1 引言	21
3.2 立体视觉测量原理	21
3.3 近景摄影测量原理	22
3.4 结构光测量原理	24
3.5 本章小结	25
第四章 计算机视觉三维测量过程与技术研究	26
4.1 引言	26
4.2 摄像机标定	26
4.2.1 传统标定方法	27
4.2.2 主动视觉标定方法	32
4.2.3 自标定方法	33
4.3 图像特征点提取	35
4.3.1 Harris 算子	35
4.3.2 Susan 算子	36
4.4 图像匹配	38
4.4.1 初始匹配	39
4.4.2 错误校正	40
4.5 实验及结果	42
4.5.1 硬件设备	42

4.5.2 软件环境	42
4.5.3 实验过程与数据	43
4.5.4 实验分析总结	47
4.6 本章小结	47
第五章 总结与展望	49
参考文献	50
致 谢	55
作者攻读硕士学位期间的发表论文情况	56

第一章 绪论

1.1 引言

计算机视觉 (computer vision) 是研究用计算机来模拟人和生物的视觉系统功能的技术学科^[1]。它是一门综合性的学科，其中包括计算机科学和工程、信号处理、物理学、应用数学和统计学，神经生理学和认知科学等。计算机视觉的目标是让计算机能够感知周围视觉世界，了解它的空间组成和变化规律，包括传感、抽象、判断、识别、理解等处理过程^[1]。

David Marr (1945-80) 是英国心理学家，他将心理学、人工智能和神经生理学的结果结合起来，对视觉的研究做出了重要贡献，他是计算视觉的奠基人。Marr 从 1973 到 1977 年在 MIT 人工智能实验室从事研究工作，并于 1982 年发表了 Vision (中译本：视觉计算理论 1988 姚国正等) 一书，该书旨在建立一个研究视觉的新框架。

Marr 认为视觉是一个信息处理的过程，目的是从图像中建立关于事物的形状与位置的描述。研究的是如何从一幅或多幅二维图像来动态的获取对对应场景的理解。重建三维景物或根据三维景物的几何形状与物理性质来识别三维景物，是计算机视觉的主要研究内容。视觉的研究可以分为 3 个层次：计算理论、表达与算法、硬件实现。将视觉信息分为三个阶段，即图像低层处理、中层处理（空间表达与建模阶段）和高层分析，其中图像低层处理主要完成图像获取，图像预处理（包括图像滤波、增强、矫正等）和抽取要素图；中层处理主要完成恢复其 2.5 维结构，形状与位置并进行建模与表达；高层分析主要完成识别、分析、理解、描述等任务^[2]。

Marr 的视觉系统框架存在以下特性^[3]，1) 没有考虑视觉中的选择性和整体性；2) 不确定性和多义性；3) 计算量大。虽然 Marr 的视觉系统框架存在很多缺陷，但 20 多年来一直在计算机视觉领域占主导地位，并在许多领域得到应用。在工业自动化方面，如工件的校验和质量控制，机器人导航，机器人的工件获取和安放、测量；在人机交互方面，如人脸的检测、跟踪、识别、建模和动画，人体检测和跟踪，手势识别，事件的检测和识别，视觉监控等等。

二十世纪 90 年代以来，基于几何的计算机视觉得到了迅速的发展，它把复杂的数学(尤其是射影几何等)引入到计算机视觉中，形成了所谓的“多视图几何”^[4]。这使得 10 年前尚未解决并经常被认为无法解决的一些问题已经有了漂亮的结果，其中一个显著的成绩是在不要求预先标定摄像机内参数(并且允许摄像机的内参数在摄像过程中变化)的条件下，就能直接由图像恢复三维物体的形状、

摄像机运动以及摄像机的内参数。它为更有效更实用的重构方法提供了理论基础。它不仅免去了用特定的标定物来标定摄像机的烦琐工作，并扩展了其应用范围。

人体体表测量是医学、法医学领域的重要课题之一。医疗诊断中需要测量烧伤区域面积、瘢痕面积等作为临床诊断的依据；临床法医进行体表损伤鉴定时，需要精确测量体表损伤区域的面积、长度以及关节的活动角度等作为鉴定的重要指标。传统测量方法一般采用手工直接在体表估测或拍照后采用九宫格辅助测量，具有测量不准确、操作不方便、难以重复测量等缺陷。一些基于损伤部位图片的测量方法，比如曲面坐标网格法^[6]，通过贴在目标区域上紧贴体表曲面的两条相互垂直的纸质尺标构造一个局部曲面坐标，并根据此曲率将区域划分成小网格进行近似测量，此方法存在操作不便，测量不精确，复杂表面无法准确测量等缺陷。在国内，一些研究者在基于计算机视觉的人体三维测量领域进行了一些研究，如基于头部不同视角的实拍照片重建头部三维模型^[7]，基于双立体同步摄影的面部三维重建与量测^[8]，网格投影式立体视觉三维表面重建^[9]，中科院自动化所模式识别国家重点实验室采用平面激光的主动三维信息获取研制成功三维面像采集与重建系统^[10]，已经应用于口腔医学中等。国外的三维测量技术发展时间较长，其应用范围较广泛，具有许多先进成熟的测量系统，如德国的GOM公司应用于工业领域的Atos三维测量系统、Tritop测量三维测量系统等。国外也有人通过三维激光扫描仪进行体表重建测量，此方法测量精度较高，但存在设备昂贵、操作不方便等缺点。

随着计算机视觉、光学的理论和技术的飞速发展，近年来在工业、医检等领域出现越来越多的基于计算机视觉的三维测量方法能够对目标区域进行三维重建，并实现精确的测量。

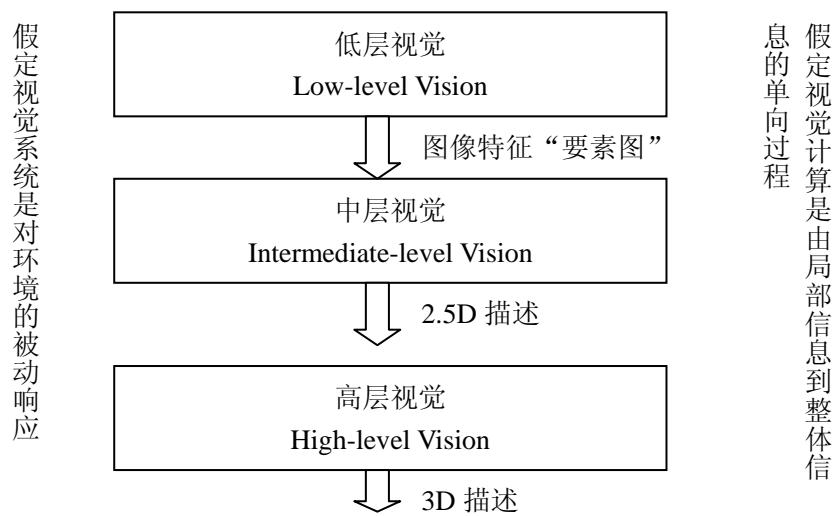


图 1-1 Marr 视觉系统框架图

1.2 三维测量技术综述

1.2.1 三维测量技术发展综述

三维测量因其实用性和重要性一直是测量领域研究的重要方向，从上世纪五十年代到至今，人们不断地研究三维测量的理论和新技术，提出了许多新方法，到现在已经形成了一系列的方法，为便于研究，对三维测量的方法进行分类^[3]，如下图所示：

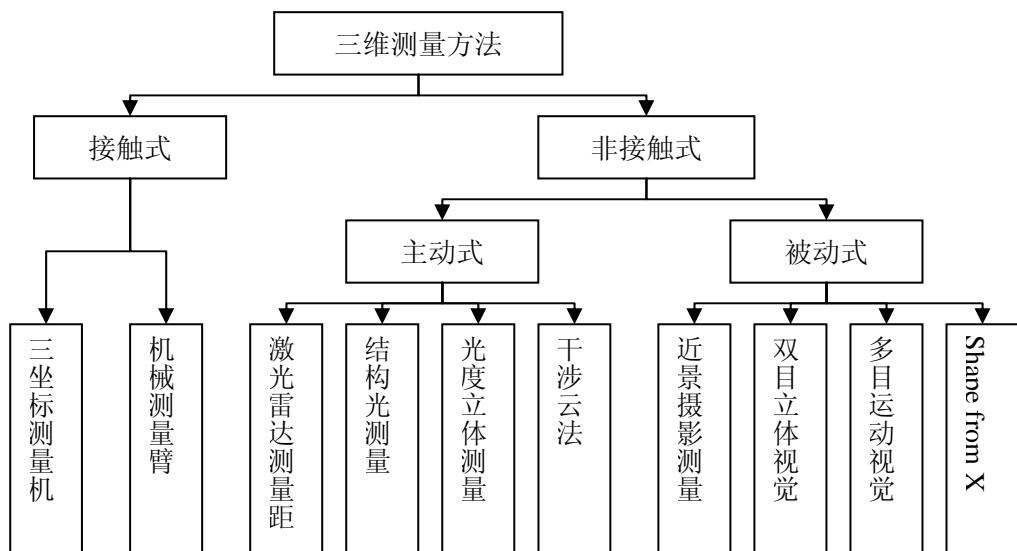


图 1-2 三维测量方法分类图

早期的三维测量一般采用接触式方法，如五十年代末出现的三坐标测量机 (Coordinate Measurement Machine, CMM)，至今仍在使用，它通过探针测头沿坐标轴方向精确移动被测对象来测量接触点的三维坐标。该方法的测量精度高，但存在价格昂贵，测量速度慢，在被测物体表面形状复杂的情况下，规划测量路径复杂困难等局限性，且对复杂物体测量和数字化建模困难。机械测量臂沿用 CMM 的测量原理并进行改进，将坐标轴探针的伺服驱动机构改进为可精确定位的多关节随动式机械臂，由人牵引装有探针的机械臂在被测物体表面滑动、扫描，实时测量机械臂各关节的转动角度可以获得被测物体的三维坐标。该方法能够克服 CMM 系统的复杂的问题，但其测量速度不能够满足快速测量的要求^[11]。

接触式测量的各种局限性，促使人们研究采用激光、超声波等非接触方式进行三维测量。测量设备主动向被测物体发射信号，通过检测返回信号来计算被测量物体的三维空间坐标。这类方法具有测量精度高的特性，但存在速度慢，对被

测物体表面和测量环境存在限制的缺点。

三维测量最具前景的是基于光学与计算机视觉理论的方法^[12]。1965年Robert发表论文“三维物体的机器感知”，提出用计算机视觉技术获得物体三维信息的可能性。20世纪70年代中期Marr提出计算机视觉系统框架，完成计算机视觉的理论体系。在此基础上三维测量具备了更加坚实的理论基础，并出现许多实用的三维测量技术和方法。基于光学和视觉理论的三维测量可分为两大类，即被动式和主动式。被动式包括传统的Shape from X(X指Shading、Stereo、Texture、Motion和Contours等)，和双目、多目视觉法^{[1][13]}等。由于视觉问题是成像过程的逆问题，要实现测量必须恢复在成像过程中丢失的深度信息^{[1][13][14]}，在成像过程中场景、对象的材质、光照、距离等在图像上仅由平面图像的像素灰度（或颜色）值反映，由唯一的灰度值反向求出多个相关信息是一个病态问题。在多目视觉中还需要解决对应点匹配问题，算法复杂、处理耗时长，难以得到满意的结果。受计算机运算速度和图像获取精确度的制约，此类方法在近年发展不大。主动式是通过测量系统主动在被测量对象表面施加干涉，包括激光、结构光等，再由通过被物体表面调制的光信息测量出物体表面的空间三维坐标。结构光包括点光源、线光源、平行光条和编码光^{[15][16]}等。这类方法的特点是算法复杂度降低，扫描速度快，无需接触被测物体表面。但由于受光电器件性能的限制，除激光扫描外三维测量外，其它方法没有得到大的推广。

二十世纪80年代以来，随着计算机处理能力快速提高、光电器件设备的性能大幅提高，同时，计算机图像处理和计算机视觉的理论和技术深入发展，基于光学、数字图像处理和计算机视觉的三维测量技术得到了很大的发展。一些发达国家研制出许多的测量新技术和应用系统，其应用领域包括医学、影像、文物、雕塑等，如人体头部、全身扫描，机器零件检测，计算机视觉导航，三维重建等。此阶段的特点是速度快、精度高、非接触。

二十世纪90年代以来，除继续发展测量新技术，提高测量精度外，还在测量的应用方面进行了扩展。测量的范围开始由简单表面、小件物体发展到复杂曲面、大型曲面的测量，在汽车工业、飞机制造等领域开始应用三维测量技术。

近年来，国内在三维测量领域也进行了研究，并在立体视觉、结构光测量、多目视觉、近景摄像测量等方面取得了一些成果。如三维面相采集与重建系统^[17]，汽车大型覆盖件三维测量^[3]，车身三维视觉检测^[18]，焊缝跟踪^[19]，堆料体积测量^[20]，人体运动跟踪^{[21][22][23]}等，应用范围包括医学、工业、交通、文物等诸多领域。由于起步较晚，跟发达国家相比，我国的三维测量技术和应用还是存在较大差距。

1.2.2 三维测量方法综述

上一节综述了三维测量技术的发展情况，根据三维测量的分类情况来看，到目前为止发展了很多实用的方法，下面就三种常用方法进行说明。

1) 近景摄影测量

摄影测量 (photogrammetry) 是以透视几何理论为基础，利用拍摄的图像来计算出于三维空间中被测物体几何参数的一种测量手段^{[13][14]}。

摄影测量按摄影距离的不同可分为遥测、远景摄影测量、近景摄影测量和显微 (超近景) 摄影测量^[13]。近景摄影测量从航空摄影测量发展而来，航空摄影测量的拍摄距离从几公里至几十公里，而拍摄距离为几米至几百米的摄影测量，被称为近景摄影测量。其测量原理与航空摄影测量相同，是通过摄影得到的图像对场景或物体的几何参数和其它参数进行测量。

近景摄影测量一般使用高精度摄像机从多个视角对被测物体进行拍照，空间任一点 P 在每个视角图像中对应唯一的像点 P_i (i 表示第 i 个视角图像)，利用不同位置拍摄时摄像机光心与像点构成的射线相关的关系，可以通过解线性方程组求得 P 点的空间三维坐标 (X_w, Y_w, Z_w) 。为求解 P 点的空间三维坐标，需要预先对摄像机进行标定计算出摄像机的内外参数，并求解出摄像机在各个方位的相对位姿。

由于成像技术的限制，摄影测量主要应用领域是航空遥测和大地远景测量，但近年来随着成像技术的发展以及图像分辨率和精度的不断提高，摄影测量技术得到了快速发展，特别是近景摄影测量法受到人们的重视，已应用到许多领域^{[3][13][14][24][25][26][27]}。近景摄影测量与多目立体运动视觉测量非常相似，但算法上不同，且近景摄影测量所用摄像机一般为高精度专业摄像机。

2) 立体视觉测量

二十世纪六十年代 MIT 的 Robert 将二维图像的分析推广到三维场景，标志着立体视觉技术的开始。二十世纪七十年代末 Marr 的视觉计算理论对立体视觉的发展产生了巨大的影响。近年来立体视觉技术得到了迅速的发展。

立体视觉 (stereo vision) 建立在计算机视觉基础上，模拟人类双眼处理信息方式的三维测量方法。通过两台相对位置固定的摄像机，从不同角度同时获得同一场景的两幅不同图像，通过计算空间点在两幅图像中的像差来获得点的三维坐标。其原理如图 1-3 所示，空间任意点 J 通过成像后在两幅图像中分别对应 $J_1(X_1, Y_1)$ 和 $J_2(X_2, Y_2)$ ，经过空间三维重建算法即可由 J_1 和 J_2 计算出 J 点的空间

三维坐标 (X_w, Y_w, Z_w) 。

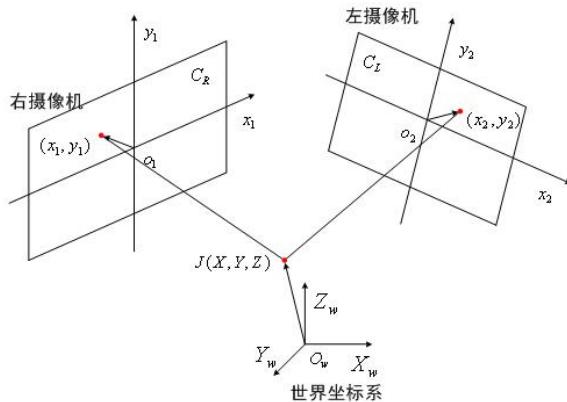


图 1-3 立体视觉基本原理图

立体视觉具有简单、灵活、可靠、使用范围广等特点，能进行非接触、自动、在线测量，在三维测量、机器人视觉、车辆自动驾驶、多自由度机械装置控制等方面具有广阔的应用前景^{[19][23][28][29]}。为解决两幅图像中对应点的匹配问题，一般采用激光扫描、编码光投影等主动方式进行测量（称主动立体视觉测量方法）。

3) 结构光测量

结构光三维测量基于光学双像机三角测量法的原理，将双像机交会中的某一个像机用事先约定好的结构投影光来代替，利用投影系统和成像系统的几何信息来进行三维形貌测量^{[3][13]}。结构光测量系统主要由结构光投影器、摄像机、图像采集及处理系统组成。结构光类型主要有：条纹光栅、正交光栅、圆形光条、点阵、交叉线、空间编码模板等。根据不同的应用对象，可设计不同的结构光形式。结构光的形式包括云纹法、单光束三维测量法和条纹结构光测量法等。

云纹法利用两组频率相近的栅线叠加产生几何干涉形成云纹条纹图，是测量面内和离面位移或变形的一种常用方法。利用这种栅线几何干涉原理，发展形成许多种云纹测量法，例如影像云纹法、投影云纹法、扫描云纹法、激光干涉云纹法、数字逻辑云纹法以及其他改进方法，使云纹等值线测量技术满足不同的实用化要求。

单光束三维表面测量系统由光源、摄像机和移动平台组成。系统光源采用倾斜激光束，摄像机为线阵 CCD 摄像机（也可使用平面光束和面阵 CCD），实际测量中，入射光保持不动，用移动装置移动被测物体来完成整个表面的测量。其测量水平分辨率取决于移动装置的最小位移量。

条纹结构光测量法采用光学投射器将光栅投影于物体表面，在表面上形成由

被测物体表面形状所调制的光栅条纹三维图像(如图 1-4 所示)。该三维条纹图像由处于另一位置的摄像机拍摄,而获得光栅条纹的二维变形条纹图像。条纹的变形程度取决于光学投射器与摄像机之间的相对位置和物体表面形廓(高度)。直观上,条纹在法线方向的位移(或偏移)与物体表面深度成比例,扭结表示了平面的变化,不连续显示了表面的物理突变或间隙。当光学投射器与摄像机之间的相对位置一定时,由变形的条纹图像便可以重现物体表面形廓,从而进行三维表面形貌测量。

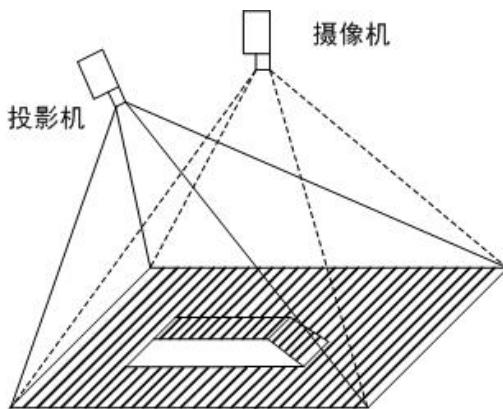


图 1-4 条纹结构光测量法示意图

结构光三维测量具有大量程、大视场、较高精度、光学条纹或结构光特征图像信息提取简单、实时性强及主动受控等特点,近年来在工业中得到了愈来愈广泛的应用。同时由于使用高能量光源照明,可以在自然光环境下工作,抗杂光干扰能力强,因此具有广阔的应用前景^{[3][15][16][30][31][32]}。

1.3 论文的主要内容及组织结构

本论文以计算机视觉和摄影测量的几何理论为基础,系统地研究人体体表三维测量的各种技术,综合分析各种技术的特点,从中选用适合解决具体应用问题的技术方法并在一些方面提出相应测量技术的改进思路和方法。

本论文共分为六章,具体内容安排如下:

第一章 根据本论文的研究内容,对相关领域的技术及其发展进行综述,并说明了本论文的主要内容和组织结构。

第二章 简要介绍了本论文研究的计算机视觉三维测量相关的理论,主要包括射影几何、摄像机成像模型、对极几何、基本矩阵等相关概念及理论体系。

第三章 研究了近景摄影测量、多目视觉测量和结构光测量这三种常用的三维测量方法的原理,和提高测量精度的相关的算法。

第四章 研究三维测量过程中一些关键技术，包括摄像机标定技术、图像特征点提取和图像匹配，并在本章的最后进行实验分析。

第五章 对本论文的研究进行总结，并对进一步研究的方向和思路进行展望。

第二章 计算机视觉三维测量理论基础

摄像机通过光学装置将三维世界的场景映射到二维的图像平面上，映射的方式由摄像机模型决定，针孔摄像机模型 (pin-hole model) 是目前广泛使用的最简单、最具体的摄像机模型，而射影几何学是描述这种摄像机成像模型的自然数学理论框架。在射影空间中，二维和三维点的坐标都用齐次坐标表示，摄像机成像模型可以方便地通过矩阵运算来描述，与摄像机成像相关的各种几何参数都可以通过矩阵运算进行计算。

本章介绍计算机视觉三维测量的理论基础，包括射影几何的相关基础理论、摄像机成像模型理论、对极几何与基本矩阵概念及理论。

2.1 射影几何

射影几何是研究射影空间在射影映射的变换群下的不变性质^[5]。摄像机的成像过程是一个射影变换 (透视或中心射影) 的过程^{[5][14]}。摄像机通过光学装置将三维世界的场景、物体映射到二维的图像平面上，物体与其图像不同，但是又有着一些共同的几何性质 (即经过映射仍保持不变的性质)。

几何学源于希腊文 *geometrein*，意思为土地测量，Euclid(约公元前 330–275) 的《原本》标志着欧氏几何的诞生。常见的旋转和平移是欧氏变换，研究在欧氏变换下保持不变的性质 (欧氏性质) 的几何，是欧氏几何。比如长度、角度、平行性等都是欧氏性质。

Pappus(约公元 3 世纪)，提出交比、对合等概念，标志着射影几何开始萌芽。Desargues(1591–1661)，引入无穷远元素，透视定理，交比、调和不变，极点、极线，创立了射影几何。射影几何是一个基础几何，而欧氏几何、双曲几何、黎曼几何等许多非欧几何都是射影几何的子几何^[5]。

2.1.1 射影空间

在 n 维欧氏空间中加入无穷远元素，并对有限元素和无穷远元素不加区分，则它们共同构成了 n 维射影空间 (projective space)。其中无穷远元素包括无穷远点、无穷远线和无穷远平面。在射影几何中，两条平行线相交于一点，这个点称为无穷远点；两个平行平面相交于一条直线，这条直线称为无穷远直线。另外规定，一条直线上只有唯一一个无穷远点，所有的一组平行线共有一个无穷远点；在一个平面上，所有的无穷远点组成一条直线，称为这个平面的无穷远直线；

三维空间中所有的无穷远点组成一个平面，称为这个空间的无穷远平面^[5]。

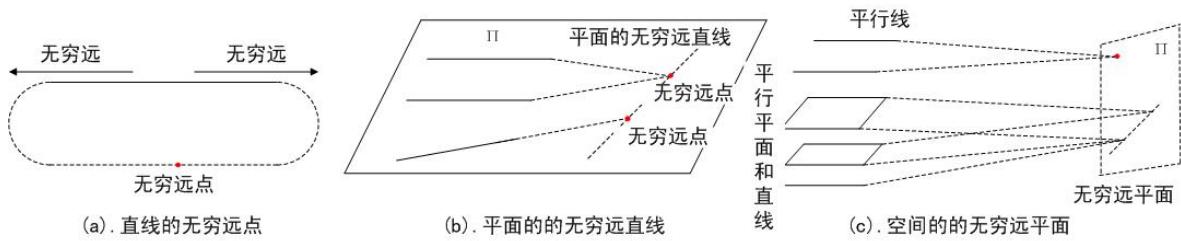


图 2-1 射影空间的无穷远元素

1 维射影空间是一条射影直线，它由欧氏直线和它的无穷点组成；2 维射影空间是一个射影平面，它由欧氏平面和它的无穷远直线组成；3 维射影空间由欧氏空间与无穷远平面组成^[5]。

在欧氏空间中建立坐标系后，点与坐标之间存在一一对应关系，但当引入无穷点以后，无穷远点无法对应欧氏坐标，为了表示无穷远点的坐标，在射影空间中引入齐次坐标。在 n 维空间中，建立欧氏坐标后，任一个有限点 M （相对于无穷点的概念）的坐标为 (m_1, m_2, \dots, m_n) ，对任意 $n+1$ 个数 $x_1, x_2, \dots, x_n, x_0$ ，如果满足：

$$\frac{x_1}{x_0} = m_1, \frac{x_2}{x_0} = m_2, \dots, \frac{x_n}{x_0} = m_n, x_0 \neq 0 \quad (2.1)$$

则 $(x_1, x_2, \dots, x_n, x_0)$ 被叫作点 M 的齐次坐标 (homogeneous coordinates)，而对应的欧氏坐标 (m_1, m_2, \dots, m_n) 被称为非齐次坐标 (inhomogeneous coordinates)。

无穷远点的齐次坐标为 $(x_1, x_2, \dots, x_n, 0)$ 。

在射影空间中直线由两点确定，直线 L 与确定直线的空间点 x_1, x_2 之间存在如下关系

$$L = x_1 \times x_2, L^T x_1 = 0, L^T x_2 = 0 \quad (2.2)$$

任意非退化二次曲线为

$$x^T C x = 0 \quad (2.3)$$

其中， $C = C^T$ ， $\text{Det}(C) \neq 0$

根据对偶原理，可使用其对偶形式表达为

$$l^T \Omega l = 0 \quad (2.4)$$

其中， Ω 为 C 的伴随矩阵，有 $\Omega = C^{-1}$

绝对二次曲线是无穷远平面上的一条二次曲线，其表达式为

$$x^2 + y^2 + z^2 = 0 \text{ 或 } X^T X = 0 \quad (2.5)$$

其中 $X = (x \ y \ z \ t)^T$, $t=0$ 。

2.1.2 射影变换

Felix Klein 在其名著 “Erlangen Program” 中提出：几何研究的是在变换群下保持不变的性质^{[4][5]}。根据几何变换其保持不变性质的程度，可分为 4 个层次：等距变换、相似变换、仿射变换和射影变换。

等距变换即欧氏变换，变换保持欧氏距离不变，只进行平移和旋转变换。其几何不变量有长度（两点间的距离）、直线的夹角、面积和体积，是刚体运动的模型。令三维点变换前的齐次坐标为 $X(x, y, z, 1)^T$ ，变换后的齐次坐标为 $X'(x', y', z', 1)^T$ 。则等距变换的数学表达式为

$$X' = H_E X = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} X \quad (2.6)$$

其中 R 是 3×3 旋转矩阵（满足 $R^T R = R R^T = I$ 的正交矩阵）， t 是三维平移矢量。

相似变换也称等形变换，是一个等距变换与一个均匀缩放的复合。其几何不变量有直线的夹角、长度的比率、面积比率和体积比率。变换的数学表达式为

$$X' = H_S X = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} X \quad (2.7)$$

其中 s 是缩放标量因子。

仿射变换是一个非奇异线性变换（非均匀缩放）与一个平移变换的复合。其几何不变量是平行关系、平行线段长度比、面积比。变换的数学表达式为

$$X' = H_A X = \begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix} X \quad (2.8)$$

其中 A 是 3×3 非奇异矩阵。

射影变换是仿射变换的推广，变换表达式为

$$X' = H_P X = \begin{bmatrix} A & t \\ V^T & v \end{bmatrix} X \quad (2.9)$$

其中矢量 $V = (v_1, v_2, v_3)^T$ 。

由以上分析知，变换层次由高到低依次为射影变换、仿射变换、相似变换和欧氏变换，高层变换可以产生比它低层的变换的所有效应^[4]。射影变换是 4 种变换中层次最高（最一般）的变换，其保持不变的几何性质最少。

下面对射影变换进行定义。

定义 2.1: 在一维射影空间中, 给定 4 点 (一维射影空间中一定共线), $x_i (i=1,2,3,4)$, 交比 (Cross) 定义为

$$Cross(x_1, x_2, x_3, x_4) = \frac{|x_1x_2||x_3x_4|}{|x_1x_3||x_2x_4|} \quad (2.10)$$

定义 2.2: 记 S_n 、 S'_n 是两个由点集构成的射影空间, P 是 S_n 到 S'_n 的映射, 若映射 P 保持以下两个性质不变, 则映射 P 被称为一个 n 维射影变换。

- 1) 共点, 共线, 接触的阶 (相交: 1 阶接触; 相切: 2 阶接触; 拐点: 与直线的 3 阶接触; 切线不连续性和歧点)。
- 2) 共线 4 点的交比。

若 S_n 、 S'_n 是同一个空间则 P 是同一空间内的射影变换。

摄像机的成像过程是一个从 3 维空间到 2 维空间的退化的射影变换^{[3][4]}。

2.2 摄像机成像模型

摄像机成像过程是将空间三维位置信息映射到二维图像平面, 这个映射关系就是摄像机成像模型^{[3][13][14]}。

2.2.1 预备知识

1) 坐标系

摄像机的成像过程是将空间三维位置信息映射到二维图像平面, 为了分析和描述成像过程中的一系列变换, 需要为成像模型建立完备的坐标系统。

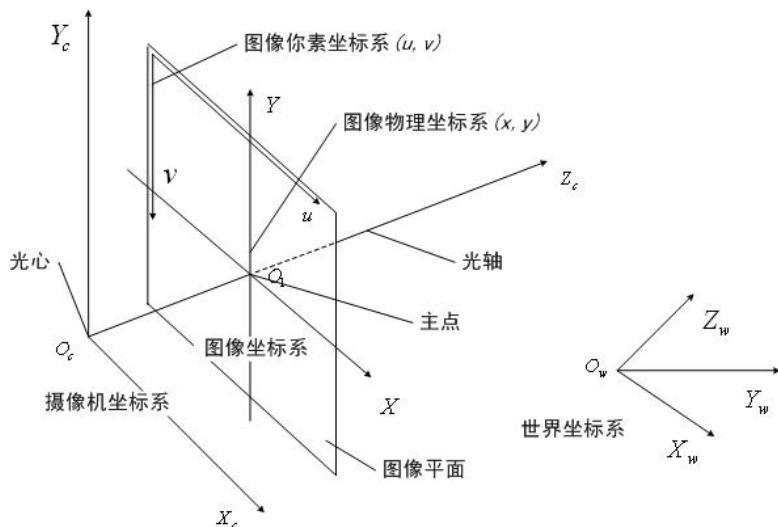


图 2-2 摄像机成像模型中的坐标系

如图 2-2 所示, 在摄像机成像模型中有三大坐标系: 世界坐标系、摄像机坐标系和图像坐标系, 其中图像坐标系又包括图像物理坐标系和图像像素坐标系。

世界坐标系 (world coordinates) 是描述真实场景中物体方位的坐标系。真实场景是三维空间, 因此世界坐标系是三维坐标, 由交于坐标原点的相互垂直的三条坐标轴构成, 坐标轴的方向可以根据实际情况设定。世界坐标系的原点和坐标轴分别记为: O_w 、 X_w 、 Y_w 、 Z_w 。

摄像机坐标系 (camera coordinates) 是描述摄像机成像系统的三维坐标系。它的原点位于摄像机的光心, 以主轴方向为坐标系的 Z 轴, X 、 Y 轴方向为摄像机的水平和垂直方向。摄像机坐标系的原点和坐标轴分别记为: O_c 、 X_c 、 Y_c 、 Z_c 。

图像坐标系 (image coordinates) 是二维坐标系, 坐标平面位于摄像机图像平面上, 它包括图像物理坐标系和图像像素坐标系。其中图像物理坐标系的原点位于主点 (一般是图像平面的中心), X 、 Y 轴方向为水平向右和垂直向上。图像像素坐标系 (也称数字化图像坐标系) 是图像数字化后描述像素位置的坐标系, 它的原点位于图像 CCD 阵列的左上角, u 为水平轴, 方向向右; v 为垂直轴, 方向向下。

2) 成像过程

摄像机成像过程是将空间三维位置信息映射到二维图像平面上。物体在三维世界中的位置由世界坐标系描述, 而最终成像在摄像机的 CCD 上, 由图像像素 (数字化) 坐标系描述, 这其中经过了多次变换, 图 2-3 描述了摄像机成像过程的变换处理过程。

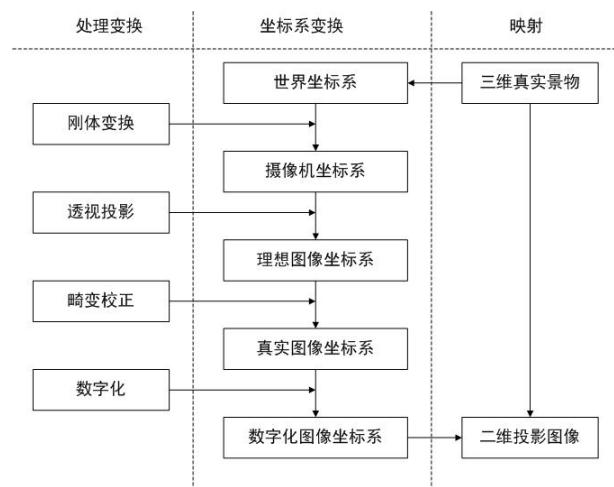


图 2-3 摄像机成像过程

对于拍摄者而言，摄像机完成了由三维景物到二维图像的一个投影变换，而对于摄像机而言，成像的过程是将真实景物的世界坐标变换到摄像机数字化图像坐标（即相机的CCD阵列位置和颜色深度），而在此过程中，需要经过刚体变换、透视投影变换、畸变校正和数字化4种变换，从世界坐标系开始经过摄像机坐标系、理想图像坐标系、真实图像坐标系，最终变换到数字化图像坐标系，得到二维投影数字化图像。

2.2.2 线性摄像机模型

在不考虑畸变的理想情形下，摄像机的成像过程由一系列的线性变换组成，是一个线性变换模型。本节主要研究摄像机的线性模型，下一节将讨论畸变因素对摄像机模型的影响。

摄像机成像的光学原理是凸透镜成像原理，其成像公式为

$$\frac{1}{f} = \frac{1}{m} + \frac{1}{n} \quad (2.11)$$

其中， f 为凸透镜的焦距， m 为像距， n 为物距。

一般情况下，由于 $n \gg f$ ，因此 $m \approx f$ ，所以透镜成像模型可以近似用小孔成像模型代替。

在计算机视觉领域中，摄像机成像模型一般都理想化为针孔模型（pin-hole model）。此处需要明确的是，针孔模型和透镜成像模型的焦距概念不同，前者的焦距等于像距，而后者不等。

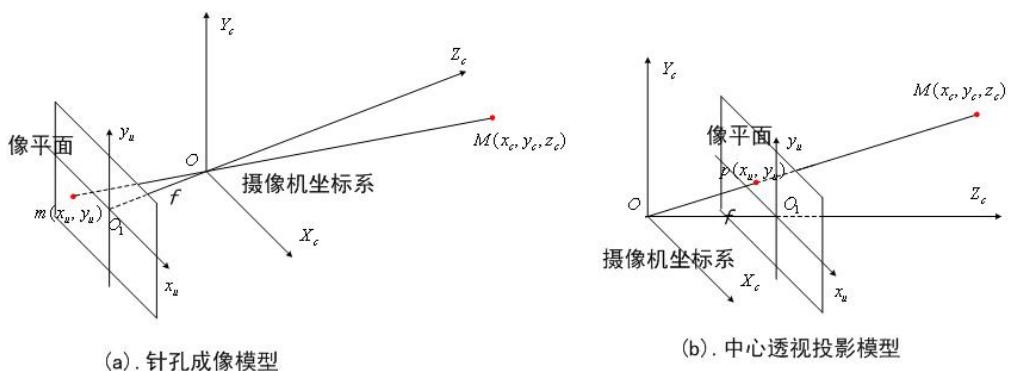


图 2-4 成像模型原理图

在上节中，摄像机的成像过程需要经过4次变换，在不考虑畸变非线性因素的情形下，空间一点 $P(x_w, y_w, z_w)$ 要经过3次变换最终在CCD像平面上成像。

1) 世界坐标系到摄像机坐标系的刚体变换

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (2.12)$$

式中 $(x_c, y_c, z_c, 1)^T$ 为 P 点在摄像机坐标系下的齐次坐标, R 是两坐标系之间的 3×3 正交旋转矩阵, t 是两坐标系之间的 3×1 平移矢量。

2) 摄像机坐标系到图像坐标系的投影变换

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} \quad (2.13)$$

式中 f 为像平面与光心之间的距离, 即等效焦距, $(x, y, 1)^T$ 为 P 的像点在图像坐标系下的齐次坐标。

3) 图像坐标系到像素坐标系的数字化变换

设像素在图像物理坐标轴上的尺寸为 dx, dy , 像素平面两坐标轴之间的夹角为 θ , 图像物理坐标系的原点在像素坐标系中的位置为 (u_0, v_0) , 如下图所示。

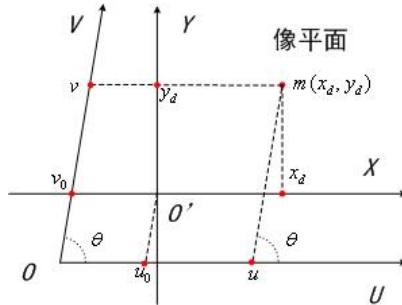


图 2-5 图像数字化坐标变换关系

则此变换为一个仿射变换, 变换表达式为

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_u & -f_u \operatorname{Cot} \theta & u_0 \\ 0 & f_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2.14)$$

其中, $f_u = \frac{1}{dx}$, $f_v = \frac{1}{dy}$ 。若考虑 $\theta = 90^\circ$ 则 (2.14) 式退化为

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2.15)$$

4) 世界坐标系到图像像素坐标系变换

综合上述变换可得到

$$\begin{aligned} z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} &= \begin{pmatrix} f_u & -f_u \operatorname{Cot}\theta & u_0 \\ 0 & f_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} ff_u & -f_u \operatorname{Cot}\theta & u_0 \\ 0 & ff_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} = K \cdot M \cdot X_w = P \cdot X_w \end{aligned} \quad 2.16)$$

其中, P 为 3×4 的投影矩阵; K 只与摄像机的内部参数相关, 称之为摄像机内部参数矩阵, 共包含 6 个变量; M 由摄像机在世界坐标系中的相对位姿确定, 称之为摄像机外部参数矩阵, 共包含与平移、旋转相关的 6 个变量。

式 (2.16) 即摄像机在不考虑镜头畸变效应的线性模型。

若 $\theta = 0$ 则式 (2.16) 可简化为

$$z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} ff_u & 0 & u_0 \\ 0 & ff_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} = K \cdot M \cdot X_w = P \cdot X_w \quad (2.17)$$

由式 (2.16) 或式 (2.17) 描述的摄像机模型知, 若确定 K 和 M 即摄像机的内外部参数, 就可以确定摄像机的投影参数。计算摄像机内外部参数矩阵 K 和 M 的过程被称之为摄像机标定 (Camera calibration), 摄像机标定是所有三维测量方法都需要完成的重要环节, 在第四章中将专门研究摄像机标定的方法。

2.2.3 摄像机畸变

由于透镜加工和安装等多方面因素影响, 摄像机的线性模型不能准确地描述其成像几何关系, 尤其是镜头质量不佳或在使用视角较大的广角镜头时, 在远离

图像中心的边缘区域会形成较大的畸变^[3]。

Tsai 1986 年提出的摄像机非线性模型认为镜头的畸变误差主要是径向轴对称畸变^[33]，Weng 1992 年在 Tsai 的基础上提出摄像机非线性误差的修正模型，Weng 的模型中提出了切向畸变的模型，并将畸变分为三类：径向畸变（Radial distortion）、离心畸变（Decentering distortion）和薄棱镜畸变（Thin Prism distortion）^[34]。三种畸变对图像会产生径向和切向两种失真。

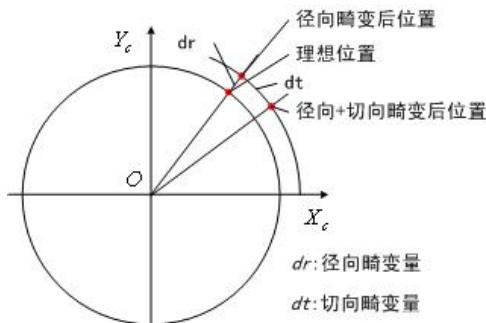


图 2-6 Weng 的摄像机畸变模型

Weng 的径向畸变与 Tsai 的径向轴对称畸变是一致的，是由于镜头在加工时的误差造成的，特点是像点位置的误差与其离光心的距离正向相关。离心畸变是由于透镜组中各透镜的光心不共线造成的，在成像过程中径向和切向都会造成畸变。薄棱镜畸变是由于镜头的设计及安装不当造成的，成像时也会在径向和切向造成畸变。

考虑摄像机畸变因素的影响，Tsai 和 Weng 分别提出了各自的非线性摄像机模型。与线性模型相比，非线性模型更接近实际情况，也更加准确，但却存在算法复杂，计算量大等缺点。

2.3 对极几何与基本矩阵

对极几何（Epipolar geometry）是同一场景在两个不同视点拍摄的两幅图像之间的射影几何关系，它与景物无关，只依赖于摄像机的内参数和相对位姿。基本矩阵（Fundamental matrix）是对极几何关系的矩阵形式的代数表示，是秩为 2 的 3×3 矩阵^[5]。

2.3.1 对极几何

图 2-7 描述了同一空间点在两个不同视点拍摄的两幅图像之间的射影几何关系，其中 $e1$ 、 $e2$ 为外极点， $L1$ 为外极线， $L2$ 为对极线， $C1C2$ 和 $e1e2$ 为基线，

平面 $C1C2$ 即 Π 为外极平面。

如图 2-7 所示, 空间任意一点 P 在左右摄像机像平面成像后, 其像点 $m1, m2$ 满足对极几何约束, 即左图像中的 $m1$ 在右图像中对应一条直线 $L2$, 同样右图像中的 $m2$ 在左图像中对应一条直线 $L1$, $m1$ 与 $m2$ 互为对应点 (或称匹配点)。左像平面中任意一点在右像平面中的对极线必然经过外极点 $e2$, 反之亦然, 对极关系具有对称性。所有的外极线都过对应的外极点, 外极点是光心连线与图象平面的交点, 对应外极线束构成一个射影变换。

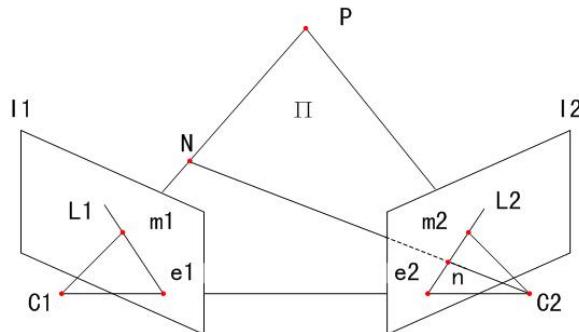


图 2-7 对极几何原理图

$I1, I2$ 为两摄像机的像平面, $C1, C2$ 为两摄像机的光心, P 为空间点, $m1, m2$ 为空间点 P 在两摄像机像平面上的像, $e1, e2$ 为两摄像机光心连线与其像平面 $I1, I2$ 的交点

当摄像机作平行于图像平面的平移或以垂直于图像平面的直线为旋转轴旋转时, 基线与图像平面的交线位于无穷远, 因此对极点为无穷远点, 而对极线为平行线。

对极几何中的对极约束关系可用在三维测量中的图像匹配、三维重建和摄像机自标定等方面。

2.3.2 基本矩阵

基本矩阵是对极几何关系的矩阵形式的代数表示。根据 2.2 节摄像机成像模型, 令 $X = [x, y, z]^T$ 为空间中任意一点, 其齐次坐标为 $\tilde{X} = [x, y, z, 1]^T$, 在像平面上的成像为 $m = [u, v]^T$, 齐次坐标为 $\tilde{m} = [u, v, 1]^T$, 则有

$$s \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = P \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (2.18)$$

即 $s \cdot \tilde{m} = P \cdot \tilde{X}$

其中, s 是比例因子, P 是投影矩阵, $P = A \cdot [R, T]$: A 为摄像机内部参数矩阵, R, T 为摄像机外部参数。该变换将摄像机坐标系下的点转换到图像坐标系下。对于 CCD 相机而言, 为简化推导, 假设 u, v 两轴之间的夹角 θ 为 90 度。则 A 可表示为

$$A = \begin{pmatrix} a_u & 0 & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.19)$$

在双目成像中, m_1, m_2 对应于空间一点 M , 由外极线几何约束可知, m_1, m_2, C_1, C_2 一定位于同一平面中, 故外极线约束也称同平面约束。设摄像机 1 到摄像机 2 有位移为 (R, T) , 由 (2.17) 式得

$$s_1 \cdot \tilde{m}_1 = A_1 \cdot [I \ 0] \cdot \begin{bmatrix} X \\ 1 \end{bmatrix} \quad (2.20)$$

$$s_2 \cdot \tilde{m}_2 = A_2 \cdot [R \ T] \cdot \begin{bmatrix} X \\ 1 \end{bmatrix} \quad (2.21)$$

其中 A_1, A_2 为两摄像机的内部参数矩阵, 由 (2.20)、(2.21) 消去比例因子 s_1, s_2 有

$$\tilde{m}_2 A_2^{-T} [T]_x R A_1^{-1} \tilde{m}_1 = 0 \quad (2.22)$$

其中 $[T]_x$ 为 T 的反对称矩阵, 由平移矢量 T 决定, 设 $[T] = (T_x, T_y, T)^T$, 则

$$[T]_x = \begin{pmatrix} 0 & -T_y & T_z \\ T_y & 0 & T_x \\ -T_z & T_x & 0 \end{pmatrix} \quad (2.23)$$

$[T]_x$ 是一个斜对称矩阵。令

$$F = A_2^{-T} [T]_x R A_1^{-1} \quad (2.24)$$

$$E = T \times R = [T]_x R \quad (2.25)$$

则有

$$\tilde{m}_2 A_2^{-T} E A_1^{-1} \tilde{m}_1 = 0 \quad (2.26)$$

$$\tilde{m}_2 F \tilde{m}_1 = 0 \quad (2.27)$$

式 (2.24) 对极线约束的数学描述, 式 (2.24) 中的矩阵 F 就是基本矩阵 (Fundamental matrix), 它与摄像机的内部参数和两相机之间的相对位姿有关。式 (2.25) 中的 E 仅与两摄像机之间的相对位姿有关, 称为本质矩阵 (Essential matrix)。

根据推导过程可知, 基本矩阵 F 具有如下性质:

- 1) 对于两幅图像中的任意一对对应点 (或称匹配点) $x \leftrightarrow x'$, 必定满足约束 $x^T F x = 0$ 。
- 2) F 为自由度为 7、秩为 2 的 3×3 矩阵。
- 3) $L_1 = F \cdot m_2$ 是对应于 m_2 的对极线, $L_2 = F \cdot m_1$ 是对应于 m_1 的对极线。另外, 如图 2-7 所示, 如果 m_1 位于极线 L_1 上, m_2 位于极线 L_2 上, m_1 和 m_2 不一定是对应点, 下述关系仍然成立:

$$n^T F m_1 = 0 \quad (2.28)$$

- 4) 两幅图像中的对极点 e_1 、 e_2 分别对应 F 的右、左零矢量, 即有

$$F \cdot e_1 = 0, \quad e_2 \cdot F = 0 \quad (2.29)$$

基本矩阵的求解算法可分为线性算法和非线性算法两类, 在线性算法中有 7 点法、8 点法、线性最小二乘法等; 非线性算法有鲁棒法、再投影法等。

2.4 本章小结

本章首先介绍了射影几何的基础理论知识, 这是摄像机成像模型和对极几何的基础; 在射影几何理论的基础上, 根据摄像测量原理推导出了摄像机成像的线性模型, 并介绍了畸变对成像模型的影响; 最后研究分析了计算机视觉中非常重要的对极几何理论原理和基本矩阵的概念。本章的内容以理论分析为主, 为后续章节的算法分析提供了理论基础。

第三章 计算机视觉三维测量方法研究

3.1 引言

人体体表三维测量是由二维图像恢复人体体表在世界坐标系中的三维面信息来完成的，在第一章中的三维测量技术综述中我们知道要测得三维信息可以采用多种方法，可以借助专用硬件设备实现的接触式测量，或者采用各种非接触式的测量。因为人体表面为柔软的表皮，不具有刚性，使用接触式测量很难完成；另外，对于在对损伤区域进行测量时，一般不能接触伤口，因此必须采用非接触式测量方法。在诸多的非接触式测量方法中，本论文主要研究基于计算机视觉的常用测量方法，即利用摄像机成像模型等即能从获取的二维图像恢复三维结构的方法^{[3][4]}。

多年来许多计算机视觉领域的学者就利用摄像机模型获得的二维图像序列恢复物体的三维结构的问题提出了很多算法。如立体视觉方法、分层重建方法、近景摄影测量方法、结构光测量方法等。对于各种方法的基本原理和特点在第一章中已经介绍，本章主要研究最常用的三种方法的原理：立体视觉测量方法、近景摄影测量方法和结构光测量方法。

3.2 立体视觉测量原理

立体视觉三维测量是在已知多个摄像机（通常是两个）相对位置的情况下，由多幅图像恢复被测物体三维形状。下面说明两摄像机组成的立体视觉测量系统的原理和基本算法。

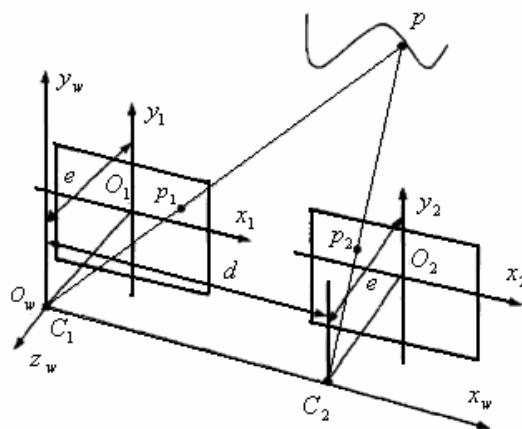


图 3-1 立体视觉三维测量原理

用 C_1 与 C_2 两个摄像机同时观察 P 点, C_1p_1 与 C_2p_2 是摄像机光心与对应像点 p_1 、 p_2 组成的射线, 通过特征点匹配方法确定 P 点在两摄像机图像上的对应关系 $p_1 \leftrightarrow p_2$, 由于空间点 P 既位于 C_1p_1 上, 又位于 C_2p_2 上, P 点是 C_1p_1 、 C_2p_2 两条直线的交点, 通过求解两条直线的交点就可以得到 P 点三维坐标。两摄像机光轴平行的立体视觉三维测量原理如图 3-1 所示。设两摄像机的投影矩阵分别为 P_1 、 P_2 , 因为立体视觉中两摄像机的位置一般是固定的, 可以预先对摄像机进行精确标定, 从而求得 P_1 、 P_2 的精确数值。因此有

$$z_{c1} \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} = P_1 \cdot \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 & m_{34}^1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (3.1)$$

$$z_{c2} \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} = P_2 \cdot \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11}^2 & m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{21}^2 & m_{22}^2 & m_{23}^2 & m_{24}^2 \\ m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (3.2)$$

由式 (3.1) 的 (3.2) 消去 Z_{c1} 和 Z_{c2} 可得关于 x_w 、 y_w 、 z_w 的四个线性方程

$$\begin{aligned} (u_1 m_{31}^1 - m_{11}^1) x_w + (u_1 m_{32}^1 - m_{12}^1) y_w + (u_1 m_{33}^1 - m_{13}^1) z_w &= m_{14}^1 - u_1 m_{34}^1 \\ (v_1 m_{31}^1 - m_{21}^1) x_w + (v_1 m_{32}^1 - m_{22}^1) y_w + (v_1 m_{33}^1 - m_{23}^1) z_w &= m_{24}^1 - v_1 m_{34}^1 \end{aligned} \quad (3.3)$$

$$\begin{aligned} (u_2 m_{31}^2 - m_{11}^2) x_w + (u_2 m_{32}^2 - m_{12}^2) y_w + (u_2 m_{33}^2 - m_{13}^2) z_w &= m_{14}^2 - u_2 m_{34}^2 \\ (v_2 m_{31}^2 - m_{21}^2) x_w + (v_2 m_{32}^2 - m_{22}^2) y_w + (v_2 m_{33}^2 - m_{23}^2) z_w &= m_{24}^2 - v_2 m_{34}^2 \end{aligned} \quad (3.4)$$

上述四个方程是三维空间的四个平面方程, 两两联立为空间的直线方程, 即 C_1p_1 、 C_2p_2 的方程, 交点为点 P 。这是视觉测量中最基本的共线方程, 说明光心、像点和物点必须在同一直线上, 几乎所有基于视觉的测量理论都是以此为基础。联立式 (3.3) 和 (3.4) 式可求出 P 点的三维坐标 (x_w, y_w, z_w) 。由于四个方程只包含三个未知量, 说明四个方程中只有三个独立方程, 只需要其中三个方程就可求出唯一解。考虑到图像的畸变和噪声, 根据上述两组方程得到的两空间直线不可能严格相交, 因此在实际应用中是通过四个方程组成的超定方程组并用最小二乘法求出 (x_w, y_w, z_w) 。

3.3 近景摄影测量原理

近景摄影测量一般使用一台高精度测量摄像机在多个不同视角拍摄被测物体图像, 令空间任一点 P 在不同位置像平面上对应唯一像点 p_i , 利用不同位置摄像机光心与像点组成的射线相交就可求出 P 点三维坐标 (x_w, y_w, z_w) 。不过近景

摄影测量在测量前一般不知道各视角的相对位置关系,在测量过程中必须先计算各视角图像坐标系之间的位姿关系,才能确定各个视角的投影矩阵,进而获取三维信息^{[24][25][26]}。

由式 (3. 1) 得到不同视角下摄像机物点与像点坐标变换方程

$$z_{ci} \begin{pmatrix} u_i^k \\ v_i^k \\ 1 \end{pmatrix} = P^k \cdot \begin{pmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11}^k & m_{12}^k & m_{13}^k & m_{14}^k \\ m_{21}^k & m_{22}^k & m_{23}^k & m_{24}^k \\ m_{31}^k & m_{32}^k & m_{33}^k & m_{34}^k \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{pmatrix} \quad (3.5)$$

其中, $(u_i^k, v_i^k, 1)^T$ 为空间任一点 p_i 在第 k 幅图像坐标系中像点的齐次坐标; $(x_w, y_w, z_w, 1)^T$ 为 p_i 点在世界坐标系中的齐次坐标; P^k 为第 k 个视角摄像机的投影矩阵, m_{ij}^k 为 P^k 的第 i 行第 j 列元素。

假设摄像机内参数通过标定已知,根据各视角图像特征点的对应关系,由点的匹配结果可恢复各视角之间的位置关系,这样投影矩阵 P^k 就可以确定。同样根据式 (3. 5) 可以得到关于 (x_w, y_w, z_w) 的共线方程

$$\begin{bmatrix} u_i^k m_{31}^k - m_{11}^k & u_i^k m_{32}^k - m_{12}^k & u_i^k m_{33}^k - m_{13}^k \\ v_i^k m_{31}^k - m_{21}^k & v_i^k m_{32}^k - m_{22}^k & v_i^k m_{33}^k - m_{23}^k \end{bmatrix} \begin{pmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \end{pmatrix} = \begin{bmatrix} m_{14}^k - u_i^k m_{34}^k \\ m_{24}^k - v_i^k m_{34}^k \end{bmatrix} \quad (3.6)$$

式 3.6 说明物点、光心和像点这三点必须在同一条直线上。这是针孔模型(中心投影)的数学表达式。

当摄像机在 N 个视角进行观察时,可以得到 $2N$ 个方程组成的超定方程组,用最小二乘法对方程组求解可得到空间点 P 的三维坐标,该算法与立体视觉算法相一致。

对近景摄影测量中空间点三维坐标计算的超定方程组求解时,一般采用最小二乘法对结果进行调整求精。捆绑调整 (Bundle Adjustment) 方法^{[3][5]}是近景摄影测量中解决多视图中空间点坐标值精度优化的最常用的方法。捆绑调整的基本算法是同步最小二乘法,它能够容忍数据的丢失并提供真正的最大似然估计;同时它允许对每一个测量值指派单个协方差并可以扩展到先验估计和摄像机参数或点的位置的约束,基本上可以称为一个理想的算法。但存在两个不足: 1) 需要提供一个良好的初值; 2) 涉及到大量的参数,可能成为一个非常大的最小化问题。

在计算机视觉三维测量领域,由于捆绑调整算法具有良好的特性,因而得到了广泛的应用,一些研究者在实际使用时通过各种方法对算法进行了优化了调整,使之更好的适应不同的应用需要。如 Kin Hong Wong 对捆绑调整进行了约束优化^[35],降低了错误率并提高了精度。

3.4 结构光测量原理

在第一章的三维测量方法综述中对结构光测量方法进行了介绍，本小节主要讨论条纹结构光测量法。

图 3-2 为条纹结构光成像模型，其中 (x_w, y_w, z_w) 为世界坐标系， (x_c, y_c, z_c) 为摄像机坐标系， (X, Y) 为图像坐标系。摄像机坐标系的原点在世界坐标系中的坐标为 $(I_x, 0, I_z)$ 。将世界坐标系下的点变换到摄像机坐标系的步骤为：首先作平移 $[-I_x, 0, -I_z]$ ，然后绕 $o_c x_c$ 轴旋转 180° ，再绕 $o_c y_c$ 轴旋转 β 角，最后再交换到图像坐标系中。

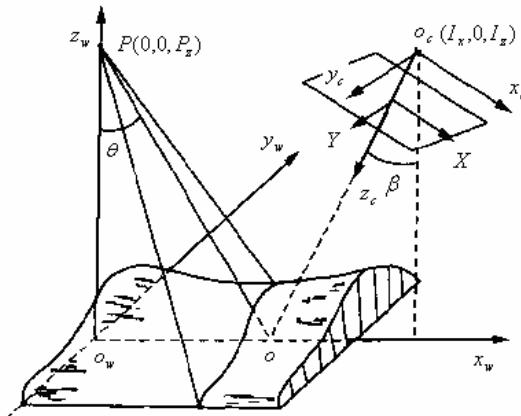


图 3-2 条纹结构光成像模型

世界坐标系与摄像机坐标系坐标变换公式为

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{pmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & -1 & 0 \\ -\sin \beta & 0 & -\cos \beta \end{pmatrix} \begin{bmatrix} x_w - I_x \\ y_w \\ z_w - I_z \end{bmatrix} \quad (3.7)$$

摄像机坐标系与图像物理坐标系的关系为

$$\begin{cases} X_u = f \frac{x_c}{z_c} \\ Y_u = f \frac{y_c}{z_c} \end{cases} \quad (3.8)$$

设结构光入射平面通过点 $(0,0,P_z)$ ，其方程为

$$z_w = -x_w \cot \theta + P_z \quad (3.9)$$

由式 (3.7)、式 (3.8) 和式 (3.9) 可以得到被测物体表面上任一点在世界坐标系下的坐标 (x_w, y_w, z_w) 与图像坐标系的变换关系，即结构光三维成像模型为

$$\begin{cases} x_w = (X_u T_1 + f T_2) / T_3 \\ y_w = \frac{Y_u}{f T_3} [(X_u T_1 + f T_2) (\sin \beta - \cos \beta \cot \theta) - T_1 T_3] \\ z_w = -x_w \cot \theta + P_z \end{cases} \quad (3.10)$$

其中, f 为摄像机焦距, T_1 、 T_2 、 T_3 的计算见下式

$$\begin{cases} T_1 = I_x \sin \beta + I_z \cos \beta - P_z \cos \beta \\ T_2 = I_x \cos \beta - I_z \sin \beta + P_z \sin \beta \\ T_3 = x_u (\sin \beta - \cos \beta \cot \theta) + f (\cos \beta + \sin \beta \cot \theta) \end{cases} \quad (3.11)$$

结构光二维成像模型中的 P_z 、 I_x 、 I_z 、 θ 、 β 、 f 等参数的值可以通过摄像机标定来确定。

3.5 本章小结

本章详细研究了目前三维测量中的主流方法的原理, 即立体视觉摄影测量方法、近影摄影测量方法和结构光测量方法, 并简要介绍了捆绑调整算法。

第四章 计算机视觉三维测量过程与技术研究

4.1 引言

基于计算机视觉的体表三维测量，是通过体表的二维图像序列恢复体表表面实际的三维位置信息，此过程是摄像机拍摄二维图像的逆向过程。根据目前的研究，典型的基于计算机视觉三维测量包括以下过程：1) 摄像机（及相关辅助设备）的标定；2) 图像特征点提取；3) 图像物征点匹配；4) 特征点三维坐标计算；5) 三维融合；6) 纹理映射与三维显示。一方面，由于二维图像恢复三维信息是一个病态问题；另一方面，重建过程本身存在复杂性。因此要构建先进、完整的计算机视觉三维测量系统是一项复杂而艰巨的工作。

要完成计算机视觉领域的各种三维重建及测量，必须先对拍摄的摄像机进行标定。摄像机标定得到的摄像机内外部参数是后续三维测量计算过程的重要参数，标定的准确及精确程度直接影响三维重建及测量最终结果的准确和精确程度。随着计算机视觉技术的不断发展，以及计算机视觉技术的应用范围日益广泛，新的摄像机标定方法不断涌现，现在研究者们提出的摄标像机定方法数不胜数。当然，不胜枚举的标定方法各有特点，以适应各种不同的三维测量场合。

计算机视觉领域里的一个瓶颈问题是在同一场景的图像间建立对应点的匹配，许多重要的计算机视觉理论与应用，如视觉测量、三维重建、监控、运动分析等都是在假设图像匹配问题已解决的前提下展开的^[3]。因此，如何实现图像对应点匹配对于进一步研究视觉计算问题具有重要的意义。而在图像匹配之前应该先确定各幅图像自身的特征点一即提取图像的特征点。图像特征点提取属于数字图像处理的问题，本章第三节中研究图像特征点提取的方法。

4.2 摄像机标定

基于计算机视觉的三维测量是在研究成像过程先验知识的基础上，建立摄像机成像模型，采用各种方法恢复出物体丢失的空间深度信息。根据本论文第二章的摄像机成像机模型，摄像机在成像时除受拍摄场景（外界光照强度与方向、物体表面反射能力）的影响外，还与摄像机的内外部参数（包括摄像机焦距、CCD阵列间距、拍摄时摄像机的位姿等）相关。为了逆向成像过程实现由二维图像完成三维重建，必须预先建立摄像机的成像模型（简称摄像机模型）并计算出摄像机的内外部参数，前者称为摄像机建模，后者称为摄像机标定^{[3][13]}。

根据摄像机成像模型和射影几何理论，摄像机标定的方法分为三类：传统标

定方法、主动视觉标定方法和自标定方法。本节以标定方法分类为线索详细研究各种标定方法。

4.2.1 传统标定方法

传统标定方法使用预先精细加工的标定块，通过建立实际标定块上已知三维点坐标与拍摄的标定块图像点之间的对应关系，来计算摄像机的内外部参数的摄像机标定法被称之为传统标定方法。传统标定方法具有适用于任意摄像机模型，标定精度高的特点，但存在标定过程复杂，标定块精度要求高，不适用于某些无法使用标定块的场合的缺点。经过长期不断的理论研究和实践应用，传统标定方法是成为目前使用最广泛的摄像机标定方法^[13]，标定块的使用也是不断创新，有传统基于正交面的体块^{[33][36]}、球形标定体^[37]、基于线的标定块^[38]、有基于平面的标定板^[39]。基于平面的又包括方格标定板^[40]，棋盘标定板^[41]、圆点标定板^[42]、同心圆标定板^{[43][44]}等。

1) 直接线性变换方法

Abdal-Aziz 和 Karara 于 20 世纪 70 年代初提出了直接线性变换 (Direct Linear Transformation, DLT) 摄像机定标方法，他们从摄影测量学的角度深入的研究了像机图像和环境物体之间的关系，建立了摄像机成像几何的线性模型，这种线性模型参数的估计完全可以由线性方程组的求解来实现。

由式 (2.16) 知， P 是 3×4 的投影矩阵， z_c 是未知的缩放因子。令 $P_{3 \times 4} = (P_{ij})$ ，由式 (2.16) 消去缩放因 z_c 得方程组

$$\begin{cases} p_{11}X_w + p_{12}Y_w + p_{13}Z_w + p_{14} - p_{31}uX_w - p_{32}uY_w - p_{33}uZ_w - p_{34}u = 0 \\ p_{21}X_w + p_{22}Y_w + p_{23}Z_w + p_{24} - p_{31}vX_w - p_{32}vY_w - p_{33}vZ_w - p_{34}v = 0 \end{cases} \quad (4.1)$$

当已知 N 个空间点和对应的图像上的点时，可以得到一个含有 $2N$ 个方程的方程组

$$A_{2N \times 12} L_{12 \times 1} = 0 \quad (4.2)$$

其中 $L = [P_{11}, P_{12}, P_{13}, P_{14}, P_{21}, P_{22}, P_{23}, P_{24}, P_{31}, P_{32}, P_{33}, P_{34}]^T$ ，即包含投影矩阵元素的向量。

根据此模型，摄像机标定的任务就是通过 (u_k, v_k) 与 (x_{wk}, y_{wk}, z_{wk}) 之间的投影关系找到合适的 L 使得 $\|AL\|$ 最小，即 $\min_L \|AL\|$ 。另外，给出约束条件

$$p_{34} = 1, \quad p_{31}^2 + p_{32}^2 + p_{33}^2 = 1$$

$$L = -(C^T C)^{-1} C^T B$$

其中, L' 为 L 的前 11 列构成的向量, C 为 A 的前 11 列构成的矩阵, B 为 A 的第 12 列构成的向量。

在上面约束条件下可以解出标定参数的线性方程组, 完成标定计算。

2) Tsai 的标定算法

20 世纪 80 年代中期 Tsai 提出的基于 RAC 的标定方法是计算机视觉摄像机标定方面的一项重要工作^{[13][33]}。Tsai 的标定方法是建立在他 1986 年提出摄像机的非线性模型的基础之上, 该模型认为镜头畸变误差主要是径向轴对称畸变^{[33][36]} (如图 4-1 所示)。

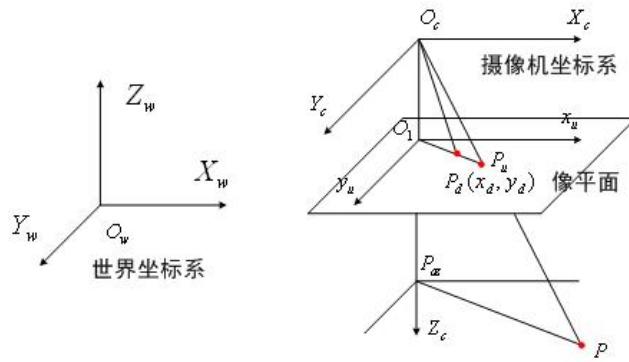


图 4-1 Tsai 的摄像机非线性模型

在 Tsai 的非线性摄像机模型中, 其畸变误差表示如下

$$\begin{cases} \delta_x = x(k_1\rho^2 + k_2\rho^4 + \dots) \\ \delta_y = y(k_1\rho^2 + k_2\rho^4 + \dots) \end{cases} \quad (4.3)$$

其中, $\rho = \sqrt{x^2 + y^2}$ 是理想像点 (x, y) 到摄像机主点的距离, k_1 、 k_2 是径向轴对称畸变误差系数。因此考虑畸变 (为方便说明只使用一阶畸变) 情况下图像坐标的变换公式如下

$$\begin{cases} u - u_0 = (x - u_0)(1 + k_1\rho^2) \\ v - v_0 = (y - v_0)(1 + k_1\rho^2) \end{cases} \quad (4.4)$$

其中 (u_0, v_0) 为图像中心 (通常把图像中心取作畸变中心和主点的坐标), (u, v) 为畸变后的数字化坐标, (x, y) 为不考虑畸变情况下像点的坐标。

考虑摄像机横向扫描与采样定时误差引起的非确定性标尺因子 s 引起的影响, 则式 (2.17) 中的 K 为

$$K = \begin{pmatrix} fs, 0, u_0 \\ 0, f, v_0 \\ 0, 0, 1 \end{pmatrix} \quad (4.5)$$

令式 (2.17) 中 $R = \begin{pmatrix} r_1 & r_4 & r_7 \\ r_2 & r_5 & r_8 \\ r_3 & r_6 & r_9 \end{pmatrix}^T$, $t = (t_1 \ t_2 \ t_3)^T$, 则有

$$\begin{cases} x = fs \frac{r_1 x_w + r_2 y_w + r_3 z_w + t_1}{r_7 x_w + r_8 y_w + r_9 z_w + t_3} + u_0 \\ y = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + t_1}{r_7 x_w + r_8 y_w + r_9 z_w + t_2} + v_0 \end{cases} \quad (4.6)$$

其中, 摄像机标定就是通过标定块上已知点空间坐标 (x_w, y_w, z_w) 与图像上提取的像点, 考虑畸变后的对应坐标 (x, y) 来计算摄像机内外参数。标定的参数主要包括: 旋转矩阵 R (三个角度未知量) 和平移向量 t (三个未知量), 等效焦距 f , 主点 (u_0, v_0) , 畸变因子 k_1 、 k_2 和非确定性标尺因子 s_x , 共计 12 个未知量。

如图 4-1 所示, P_{oz} 是 P 点在摄像机坐标系 Z 轴上的投影点, 在假定摄像机镜头畸变为径向畸变的前提下, 等效焦距 f 对 P_d 的坐标 x_d 和 y_d 的影响是同等的。空间点 P 在摄像机像平面上的理想点 P_u 与畸变点 P_d 和图像中心 O_1 三点一定共线, 且有 $P_{oz}P \parallel O_1P_d$, 这就是所谓的径向一致约束, 相应公式如下

$$\frac{x - u_0}{y - v_0} = \frac{u - u_0}{v - v_0} \quad (4.7)$$

Tsai 的标定法分为两个步骤。

步骤一: 求解像机外参数旋转矩阵 R 和 x 、 y 方向上的平移量

由式 (4.6)、(4.7) 有

$$\frac{s(r_1 x_w + r_2 y_w + r_3 z_w + t_1)}{r_4 x_w + r_5 y_w + r_6 z_w + t_2} = \frac{u - u_0}{v - v_0} \quad (4.8)$$

由至少 7 组点, 可以求得一组解

$$M_0 = (m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8) \approx (sr_1, sr_2, sr_3, st_1, r_4, r_5, r_6, t_2) \quad (4.9)$$

对 M_0 除以 $c = \sqrt{m_5^2 + m_6^2 + m_7^2}$ 则得到一组解 $(sr_1, sr_2, sr_3, st_1, r_4, r_5, r_6, t_2)$ 。

由 $r_1^2 + r_2^2 + r_3^2 = 1$ 可以求出 s , 从而 t_1 也可以被解出。

由旋转矩阵 R 的性质, 由

$$(r_7, r_8, r_9) = (r_1, r_2, r_3) \times (r_4, r_5, r_6) \text{ 或 } (r_7, r_8, r_9) = (r_4, r_5, r_6) \times (r_1, r_2, r_3)$$

可求解除出 (r_7, r_8, r_9) , 再根据 R 的性质: $\det(R) = 1$, 来选择 (r_7, r_8, r_9) 。

步骤二：求解 z 方向上的平移量、等效焦距和畸变系数

将 $k_1 = 0$ 作为初始值, 由式 (4.4) 有

$$\begin{cases} u - u_0 = x - u_0 \\ v - v_0 = y - v_0 \end{cases} \quad (4.10)$$

将 x 、 y 以及上一步骤中解出的 R 、 s 、 t_1 、 t_2 值代入式 (4.6) 有

$$\begin{cases} (u - u_0)(r_7X + r_8Y + r_9Z + t_3) = fs(r_1X + r_2Y + r_3Z + t_1) \\ (v - v_0)(r_7X + r_8Y + r_9Z + t_3) = f(r_4X + r_5Y + r_6Z + t_2) \end{cases} \quad (4.11)$$

由式 (4.11) 可以解出 f 和 t_3 。

将解出的 f 和 t_3 连同 $k_1 = 0$ 作为初始值, 对下式进行非线性优化, 可以估计出 f 、 t_3 和 k_1 的真实值。

$$\begin{cases} \frac{fs(r_1x_w + r_2y_w + r_3z_w + t_1)}{r_7x_w + r_8y_w + r_9z_w + t_3}(1 + k_1(u^2 + v^2)) = u - u_0 \\ \frac{f(r_4x_w + r_5y_w + r_6z_w + t_2)}{r_7x_w + r_8y_w + r_9z_w + t_3}(1 + k_1(u^2 + v^2)) = v - v_0 \end{cases} \quad (4.12)$$

该方法的主要优点是: 能同时求解摄像机内外参数且求解过程是线性的, 为进一步求解的非线性迭代提供了较好的初值; 考虑了摄像机坐标系三坐标轴间的正交性约束; 考虑了镜头的径向轴对称畸变, 这在多数情况下是实用的。但也存在一些局限, 如只用了平行约束条件, 没充分利用空间点与像点间的成像约束关系; 使用归一化因子求解 (r_7, r_8, r_9) 并不能保证 R 正交等; 只考虑了镜头的径向轴对称畸变, 没有考虑切向畸变, 使得畸变模型略显粗糙。

3) 基于平面模板的标定方法

张正友于1999年最早提出了一种使用平面标定模板代替标定块的方法^[40], 随后Chang Shu等也提出使用国际象棋盘作为标定物进行标定的方法^[41]。该方法只要求绘制一个具有精确定位点阵的模板, 如图4-2, 然后摄像机在不同方位(至少两个)获取平面模板的图像, 摄像机和平面模板都可以自由运动, 通过确定图像和模板上的点的匹配计算出图像和模板之间的单应性矩阵, 然后利用该单应性矩阵线性解出摄像机内参数。

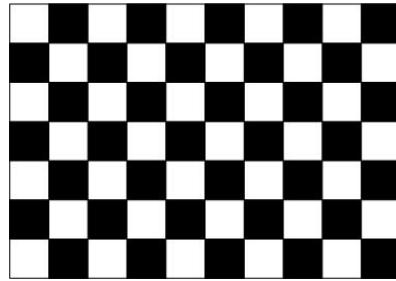


图 4-2 张正友平面标定模板

假定模板平面在世界坐标系 $z_w = 0$ 的平面上，则模板上任意一点与像点之间存在关系

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} R & t \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \quad (4.13)$$

这样标定模板平面上的点和它的像点之间建立了一个平面单应映射 H 。

$$sm = H X_w \quad (4.14)$$

$$H = [h_1 \ h_2 \ h_3] = \lambda K [r_1 \ r_2 \ t] \quad (4.15)$$

其中 λ 为尺度因子。

$$r_1 = \frac{1}{\lambda} K^{-1} h_1, \quad r_2 = \frac{1}{\lambda} K^{-1} h_2 \quad (4.16)$$

根据旋转矩阵的性质，即 $r_1^T r_2 = 0$ 和 $\|r_1\| = \|r_2\| = 1$ ，每幅图象可以获得以下两个对内参数矩阵的基本约束

$$\begin{cases} h_1^T K^{-T} K^{-1} h_2 = 0 \\ h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \end{cases} \quad (4.17)$$

由于摄像机有5个未知内参数，所以当所摄取得的图象数目大于等于3时，就可以线性唯一求解出。

张正友的平面标定方法是介于传统标定方法和自标定方法之间的一种方法。它既避免了传统方法设备要求高，操作繁琐等缺点，又较自标定方法精度高，符合办公、家庭使用的桌面视觉系统(Desktop Vision System, DVS)的标定要求。其局限性是需要确定模板上点阵的物理坐标，以及图像和模板之间的点的匹配，这给不熟悉计算机视觉的使用者带来了不便。

孟晓桥、胡占义提出圆标定方法^[42]。该方法从至少三个不同方位拍摄圆标定

模板图象,根据射影不变性计算出每幅图象上的圆环点像的坐标,得到关于内参数矩阵的至少六个方程,即可解出所有内参数。孟胡的圆模板标定方法与张正友的方法过程相似;但所用的模板不同,圆模板方法基于曲线拟合(稳定),不需要任何匹配,而张正友的方法基于模板和图像之间对应点的匹配。

吴等的平行圆标定方法利用摄像机成像的准仿射不变性,计算图像上二次曲线的交点,得到圆环点的图像进行标定。该方法和以往的基于圆的标定方法相比具有以下优点:1).从最小个数出发;2).计算圆环点图像简单;3).只需要从拟合的二次曲线出发,不需要任何匹配,不需要计算圆心;4).应用场合广泛,不仅仅限于平面的情形.也可应用基于转盘的重构。

4.2.2 主动视觉标定方法

主动视觉标定法通过控制摄像机进行某种特殊的运动,利用这种运动的特殊性计算出摄像机的内部参数。这种标定方法的优点是算法简单,能够通过线性方法求解,鲁棒性较高,但不能使用于摄像机运动未知和无法控制的场合^{[3][13][45]}。

1) 胡占义等的主动视觉标定方法

胡占义等的主动视觉标定方法^[46]可基于两种原理:基于平面单应矩阵的正交运动和基于外极点的正交运动。

基于平面单应矩阵的正交运动方法中,假定 $t^{(1)}, t^{(2)}$ 是摄像机一组正交的平移运动,存在两个单应矩阵:

$$\begin{cases} H_1 = \sigma_1(I + K \frac{t^{(1)} \vec{n}^T}{d} K^{-1}) \\ H_2 = \sigma_2(I + K \frac{t^{(2)} \vec{n}^T}{d} K^{-1}) \end{cases} \quad (4.18)$$

满足下式

$$K^{-1}(H_1^T - \sigma_1 I)K^{-T}K^{-1}(H_2 - \sigma_2 I)K = \frac{\sigma_1 \sigma_2}{d^2} \vec{n} (t^{(1)})^T t^{(2)} \vec{n}^T = 0 \quad (4.19)$$

即

$$(H_1^T - \sigma_1 I) \bar{C} (H_2 - \sigma_2 I) = 0 \quad (4.20)$$

其中 $\bar{C} = K^{-T} K^{-1}$ 。

根据式(4.20)由五组两两正交运动可以完全求解5个内参数。

基于外极点的正交运动方法中,存在如下对极关系

$$\begin{cases} e_1 \approx Kt_1 \\ e_2 \approx Kt_2 \end{cases} \quad (4.21)$$

即有

$$\begin{cases} t_1 \approx K^{-1}e_1 \\ t_2 \approx K^{-1}e_2 \end{cases} \quad (4.22)$$

从而有

$$e_2^\tau K^{-\tau} K^{-1} e_1 = t_2^\tau t_1 = 0 \quad (4.23)$$

由式 (4.23), 由五组两两正交运动变可以完全求解 5 个内参数。

这两种主动视觉标定方法与最经典的主动视觉标定方法 (马颂德的三正交运动法^[47]) 相比, 具有照相机的二正交运动比三正交运动更容易实现; 可以求解摄像机的所有 5 个内参数 (马颂德的方法可以求解 4 个内参数)。

4.2.3 自标定方法

自标定方法无需专门的标定模块, 也无明显的标定操作步骤, 仅依靠摄像机拍摄的多幅图像之间存在的对应关系进行标定。其优点是仅需要建立图像点之间的对应, 灵活性强, 潜在的应用范围广; 但存在算法非线性, 不够稳定, 鲁棒性不高的不足^{[3][13][32][41][48][49][50]}。

自标定方法有基于 Kruppa 方程和基于绝对二次曲面、无穷远平面等方法。本节主要讨论基于 Kruppa 方程的自标定方法。

1) 预备知识

所谓的自标定, 是在假定图象点之间的对应关系已经确定, 认为在拍摄不同图象时, 摄像机的内参数没有发生变化的情形下, 求解摄像机的内参数矩阵 K 。

$$K = \begin{pmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.24)$$

给定向量 $a = (a_1 \ a_2 \ a_3)$, 其对应的反对称矩阵为

$$[a]_\times = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \quad (4.25)$$

对应任意的向量 b , 有 $a \times b = [a]_\times b$ 。

2) 基于 Kruppa 方程的自标定方法

在摄像机自标定中, 最经典的方法是求解所谓的 Kruppa 方程的方法, 它是由 Faugeras 等最先引入到计算机视觉中^[51], 并且被看作是第一个自标定方法。通过绝对二次曲线建立关于摄像机内参数矩阵的约束方程, 称之为 Kruppa 方程, 它是描述二次曲线的对极切线的对应的代数表示。

在中心投影下, 如果射影平面与空间曲线相切, 则射影平面与图象平面的交线必与空间曲线在图象平面上的投影曲线相切。根据射影变换, 绝对二次曲线投影到像平面上的变换关系为

$$m \approx K(R \quad T) \begin{pmatrix} X \\ 0 \end{pmatrix} \quad (4.26)$$

由上式有

$$X \approx R^T K^{-1} m \quad (4.27)$$

由绝对二次曲线的定义式 (2.5) 有

$$m^T K^{-T} K^{-1} m = 0 \quad (4.28)$$

由上式知 K 可以通过 Cholesky 分解唯一确定; 绝对二次曲线的像仅与摄像机的内参数有关, 与摄像机的运动参数无关, 摄像机自标定的参考标定物一般选择绝对二次曲线。

由左右图像中的绝对曲线

$$\begin{cases} x_l^T C x_l = 0 \\ x_r^T C x_r = 0 \end{cases} \quad (4.29)$$

其对偶表达式为

$$\begin{cases} l_l^T \omega l_l = 0 \\ l_r^T \omega l_r = 0 \end{cases} \quad (4.30)$$

由式 (2.2) 的定义知 $l_l \approx [e_1]_x x$ (其中 x 为 l_l 是任意一点), 由基本矩阵性质有 $l_r \approx Fx$, 因而有

$$\begin{cases} x^T [e_1]_x^T \omega [e_1]_x x = 0 \\ x^T F^T \omega F x = 0 \end{cases} \quad (4.31)$$

$$[e]_x^T \omega [e]_x = \lambda F^T \omega F \quad (4.32)$$

式 (4.32) 表示的方程就是 Kruppa 方程。

$$\text{令 } C = K^{-T} K^{-1} \text{ 则 } C \text{ 的对偶曲线为 } \omega = K K^T, \text{ 其中 } \omega \approx \begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \\ \omega_2 & \omega_4 & \omega_5 \\ \omega_3 & \omega_5 & \omega_6 \end{pmatrix}$$

基于 Kruppa 方程自标定方法的基本思路包括三步：1) 通过绝对二次曲线建立关于摄像机内参数矩阵的约束方程，即 Kruppa 方程；2) 求解 Kruppa 方程确定矩阵 C ；3) 通过 Cholesky 分解得到矩阵 K 。

在 Kruppa 方程中， F, e 为已知数， ω 为未知数（有 5 个独立未知变量）；每个 Kruppa 方程最多可以提供 2 个关于未知变量的独立约束，约束方程为 5 元二次方程；每对图象可以得到一个 Kruppa 方程，故至少需要 3 对图象来标定摄像机，且摄像机的内参数必须保持不变；假定内参数都在变，任意两幅图像间有两个独立的 Kruppa 方程，则 $N (N \geq 3)$ 幅图像之间有 $N(N-1)$ 个 Kruppa 方程，其中只有 $5N-9$ 个方程是独立的。

4.3 图像特征点提取

提取图像特征点 (Feature points) 是进行图像匹配 (Image register) 的前提^{[3][4]}，本节将研究图像特征点提取的问题。图像的特征点一般是图像中灰度不连续点（即灰度或颜色变化剧烈位置处的点），在图像上一般反映在物体的轮廓、边线、角点或有特殊标志的位置上^[52]。一般将图像特征点的提取方法分为两大类：1) 先提取图像中的边缘，再在边缘上找出曲率最大点（即角点），或者将边缘用多边形来拟合，然后以多边形的顶点作为特征点；2) 直接法，即通过定义算子，通过在灰度图上寻找某种算的极值来提取特征点。

4.3.1 Harris 算子

本文采用了较为常用的 Harris 算子^[53]进行实验，该算子是基于信号的特征提取算子，其特点是能够同时提取直线和特征点。Harris 算子的表达式为

$$R(x, y) = \text{Det}(\hat{C}) - k \text{Trace}^2(\hat{C}) \quad (4.33)$$

其中 $\hat{C} = \begin{bmatrix} \hat{I}_x^2 & \hat{I}_{xy} \\ \hat{I}_{xy} & \hat{I}_y^2 \end{bmatrix}$ ，式中的 \hat{I}_x 、 \hat{I}_y 、 \hat{I}_{xy} 分别是用高斯函数滤波过的 x 、 y 方向的偏导数以及二阶混合偏导； k 是常系数。

用式 (4.33) 提取特征点时，凡满足 $R(x, y)$ 大于某一阈值 T 的像素点均被认为是特征点，因此 k 值越小则 R 值越大，检测到的特征点就多。阈值 T 依赖于实

际图像的属性,如尺寸、纹理等,但由于 T 不具有直观的物理意义,其具体值难以确定,为此,采用间接取 T 的方法:我们只需确定图像中所能提取的最大可能的特征点数目 N_{\max} ,提取程序则可根据此值选择 $R(x, y)$ 值最大的若干像素点作为特征点。

采用 Harris 算子进行特征点提取,具有算法简单,稳定和优点,但它的定位精度只能达到一个像素。谢东海等在参考摄影测量中 Förstner 算子精确定位思想的基础上,改进 Harris 算子,使其可以定位到子像素,经实验证明,该方法在点特征的提取和精确定位方面都达到了较好的效果^[3]。

4.3.2 Susan 算子

Susan 算法是 Smith 于 1995 年提出的^[13],目前在数字图像处理领域得到了广泛的应用^{[54][55]}。Susan 算法包括滤波和特征提取(边缘和特征点检测)。Susan 算法采用圆形模板进行计算处理,如图 4-3 所示,在白色背景上有一个矩形的灰色区域。

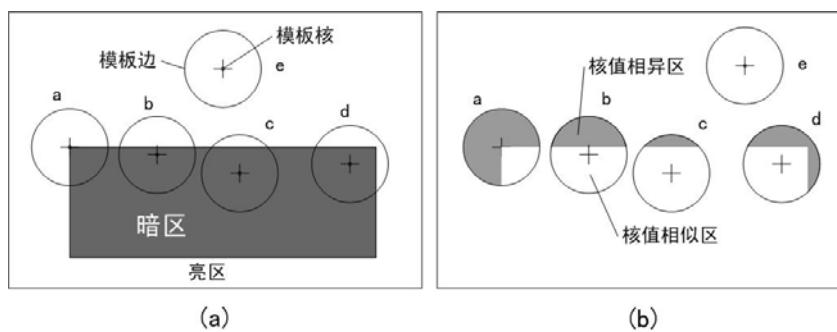


图 4-3 Susan 算法中模板与 Usan 区

当模板位于图像上的不同位置时,将模板中的各点与核心点(当前点)的灰度值用下面的相似比较函数进行计算

$$c(x, y) = \begin{cases} 1, & |I(x, y) - I(x_0, y_0)| \leq t \\ 0, & |I(x, y) - I(x_0, y_0)| > t \end{cases} \quad (4.34)$$

其中, $I(x_0, y_0)$ 和 $I(x, y)$ 分别是核心和模板中其它点的灰度值, t 是区分特征目标与背景的阈值,即相似程度阈值。阈值 t 的选取要根据图像中目标区域与背景区域的对比程度来确定。

由于图像中特征区域的内部特征(一般是灰度值)是一致的或相近的,因此对单个像素来说其特征只与周围局部区域的灰度值相近的点有关,在这个局部区域中包含着许多图像的结构信息。用相似比函数对模板中的各点与核心点的灰度

值进行比较得到与核心点灰度相近的点的集合区域称为核值相似区 (Univalve Segment Assimilating Nucleus, USAN)，即 USAN 区。图 4-3 (b) 中模板所覆盖的白色区域即为 USAN 区。

由图 4-3 可看出，在平坦区域 USAN 区域最大，如图 4-3 (e)；在平直的边缘区域 USAN 区域大小为模板的一半，如图 4-3 (b)；而在角点附近 USAN 区域最小，如图 4-3 (a)。由此可以得出 Susan 算法提取边缘和角点的基本原理为：在边缘、角点处 USAN 区域最小。因此可以由 USAN 区域的大小和矩特性来检测出边缘、角点等特征的位置和方向信息。

Susan 算子在边缘、角点检测方面与其它算子的不同在于不用对图像求导，因而其抗噪声能力强。Susan 边缘检测算法使用给定大小的模板对图像中的像素点进行运算以产生边缘初始响应，再对边缘初始响应进行处理以得到最终的边缘。采用圆形模板可以产生各向同性的边缘响应，为了提高速度也可以采用传统的 3×3 模板。不过要注意的是对模板中各点的计算均不含核心点。

在实际处理中，往往采用下面更稳定，更有效的相似比较函数

$$c(x, y) = \exp \left\{ - \left[\frac{I(x, y) - I(x_0, y_0)}{t} \right]^6 \right\} \quad (4.35)$$

Smith 指出：当理论和试验均证明指数取 6 时，错误的肯定和否定概率最小，在下图中 a, b 分别表示上述式 (4.34) 和式 (4.35) 的曲线。

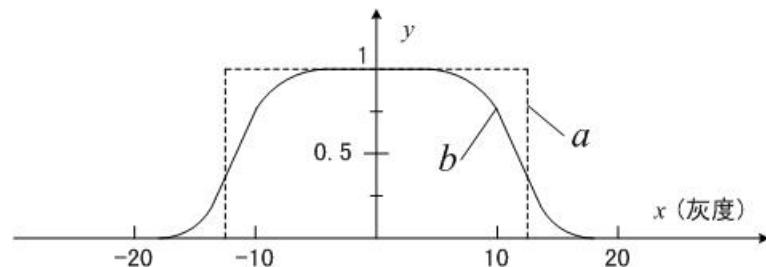


图 4-4 相似比较函数

USAN 区域的大小可由下式给出

$$n(x_0, y_0) = \sum_{(x, y) \neq (x_0, y_0)} c(x, y) \quad (4.36)$$

由上式可知对于 37 点大小的圆模板，USAN 区域的最大值 $n_{\max} = 36$ 。根据试验分析可知在实际噪音图像中如果核心点在边缘附近， n 的值一般不会大于 $3n_{\max} / 4$ 。因此定义几何阈值 $g = 3n_{\max} / 4$ ，用下式产生边缘初始响应

$$R(x_0, y_0) = \begin{cases} g - n(x_0, y_0), & n(r_0) < g \\ 0, & n(r_0) \geq g \end{cases} \quad (4.37)$$

由上式产生得到的边缘的初始响应值大小符合 Susan 原理, 即 USAN 区域越小, 初始边缘响应就越大。

SUSAN 边缘检测算法的具体步骤如下:

- 1) 用 (4.35) 式对模板中各点计算相似度;
- 2) 用 (4.36) 式计算 USAN 区域的大小;
- 3) 用 (4.37) 式产生边沿初始响应;
- 4) 用 USAN 重心及对称最长轴来确定局部边缘方向;
- 5) 在局部边缘垂直方向上取初始响应的局部极大值点的位置为边缘点, 进行细化、平滑、间断点连接、假边缘点消除和边缘分支处理等以获得单一、平滑的边缘输出。在此基础上, 还可以利用亚像素方法获取亚像素级精度的边缘信息。

由于要利用局部边缘方向来确定初始响应的局部极大值点的位置为边缘点, 下面给出局部边缘方向的确定算法: 如果 USAN 区域的重心与模板核心位置不重合, 则其局部边缘方向为重心与核心连线的垂直方向; 而对于重心与核心重合的情况, 可以利用二级中心矩法来求出 USAN 的对称最长轴方向, 即为局部边缘方向。

利用同样的思想可以进行角点的检测。具体算法与边缘检测算法类似, 但有两点需要说明: 1) 几何阈值 g 在这里取 $n_{\max} / 2$, 这恰好是理想边缘的 USAN 区大小, 而对于有噪声影响的图像边缘的 USAN 区一般都大于 g , 因此可以抑制非角点的产生; 2) 利用 USAN 区重心与核心点连接上的像素点的边缘初始响应值相近的条件来消除错误的角点。

SUSAN 算法的不足之处在于: 1) 固定阈值不适合与一般的情况; 2) 图像中噪声对边缘点提取有较大的影响。因此可以研究阈值 t 自适应选取算法。

4.4 图像匹配

图像匹配是建立同一场景下多幅图像中的对应点的关联。图像匹配是计算机视觉领域的关键问题, 因为匹配的正确性和精确度将直接影响三维重建的结果, 所以图像匹配成为计算视觉中技术中的一个瓶颈^{[3][4]}。

当图像无纹理和遮挡时, 问题变得不可解。而在实际情况下, 对应点匹配问题固有的不明确性可以使用几个约束来降低。目前已有大量文献介绍了求对应的方法, 基本思想大多是基于对极线约束、唯一性约束、相似性度量和连续性假设等^[4]。图像匹配的方法可以分为两大类: 1) 基于区域的匹配方法; 2) 特征匹配方法。

图像匹配方法一般是先使用图像匹配方法得到初始匹配集; 再以初始匹配集为基础, 通过鲁棒估计基本矩阵 F 恢复对极线约束, 并结合唯一性约束、相似性

度量和连续性假设对初始匹配集进行筛选，最后利用冗余数据进行拟合求精，最终可以得到可靠的特征点匹配集。

4.4.1 初始匹配

Zhang 的基于极线约束的匹配方法，是将两类图像匹配方法结合起来的阶段式匹配算法^[51]，目前已得到了广泛应用。该方法建立在假定匹配点周围的灰度分布相似（即灰度相似性时 Cross-correlation）的前提下，因而无法解决两幅图像视差较大的情形。一些学者针对这个缺陷提出了一些改进的方法，如 Pritchett 等用单应矩阵(Homography)取代灰度相似性和极线约束作为匹配的准则^[56]，该方法对于某些含有丰富平面信息的图像特别有效；Lhuillie 等提出了一种稠密匹配的新策略^[57]，对于纹理稠密的图像特别有效，缺陷是精度不高。

由于灰度相似性存在局限性，一些学者尝试构造能描述特征点周围的局部信息，并且在视点改变或光照强度发生变化时能尽量保持不变的特征点的“特征向量”，以取代灰度相似。如 Schmid 等的特征点一至三阶偏导联立特征向量^[58]。Tuytelaars 等的低阶分量系列^[59]。Telle 等考虑两特征点之间连线的灰度分布匹配策略^[60]。Joon 提出利用极线约束的轮廓匹配方法^[60]。

本文以 Zhang 的方法构造图像的自动匹配的初始匹配集 T 。下面为灰度相关系数的计算公式

$$s(m_1, m_2) = \frac{\sum_{i=-n}^n \sum_{j=-n}^n [I_1(u_1 + i, v_1 + j) - \bar{I}_1(u_1, v_1)][I_2(u_2 + i, v_2 + j) - \bar{I}_2(u_2, v_2)]}{(2n+1)^2 \sigma(I_1) \sigma(I_2)} \quad (4.38)$$

其中 m_1 、 m_2 分别对应两幅匹配图像中的一个特征点，其坐标分别为 (u_1, v_1) 、 (u_2, v_2) ， $\bar{I}_1(u_1, v_1)$ 、 $\bar{I}_2(u_2, v_2)$ 分别是两图像中窗口的灰度均值， $\sigma(I_1)$ 、 $\sigma(I_2)$ 是 I_1 、 I_2 的方差， $s(m_1, m_2)$ 就是在两幅图像中分别以 m_1 、 m_2 为中心的 $(2n+1) \times (2n+1)$ 方块（即图像的窗口）的灰度相关系数。当然，为了避免不必要的计算， m_1 、 m_2 之间的距离应该在预先设定的阈值范围内。

实际操作中把所有满足 $s(m_1, m_2)$ 大于阈值的点对加入到候选匹配对集合 T 中。这样得到的集合 T 通常含有以下三类误匹配：1) 同一特征点对应多个匹配；2) 不应匹配上的点对由于其灰度相关值较高而被收入 T （对于有重复图案的图像，第二类错误尤其严重）；3) 由于视差或其它原因，某些应该匹配上的点对没有被收入 T 。

4.4.2 错误校正

初始匹配中的错误可以通过对极线约束、唯一性约束、相似性度量和连续性假设等条件来校正。

1) 松弛法

利用松弛法可以在很大程度上纠正 T 中的第一类错误。这里的松弛法，是指允许 T 中的候选匹配对通过自我解散自我重新配对，使得“唯一性”和“连续性”得到最大的满足^[3]。

连续性约束以匹配强度^[51]作为参考，通过下式进行计算

$$P(m_1, m_2) = 1 - \frac{S'_M(m_1, m_2)}{S_M(m_1, m_2)} \quad (4.39)$$

其中， $S_M(m_1, m_2)$ 为匹配对 (m_1, m_2) 的匹配强度，当 (m_1, m_2) 邻域内的候选匹配对多时，则匹配强度大； $S'_M(m_1, m_2)$ 指 (m_1, m_2) 邻域内匹配强度仅次于 $S_M(m_1, m_2)$ 的最大匹配强度。因此， $P(m_1, m_2)$ 反映了 (m_1, m_2) 是唯一正确匹配的概率。

松弛过程应该兼顾到连续性和唯一性，为此构造两个数组 A_s 和 A_p ，其中 A_s 存储 T 中匹配强度 $S_M(m_1, m_2)$ 值最大的 q 对候选匹配对， A_p 存储 $P(m_1, m_2)$ 值最大的 q 对候选匹配对。每次循环时，选取同时位于 A_s 、 A_p 前 60% 的匹配对作为最后的输出，这些匹配对可以被认为在一定程度上兼顾了连续性和唯一性^[3]。

2) 随机抽样一致算法

采用对极线约束能够纠正初始匹配集 T 中的第二类错误，然而对极约束需要先计算出基本矩阵 F 。求解基本矩阵 F 的方法有很多，如张正友提出的在射影空间中变换图像点的参数化方法^[63]，能够将原来估计基本矩阵需要的 36 个参数简化成只需要 1 个参数的非线性优化；Stefan Lehmann 等提出的基本矩阵鲁棒确定法^[64]，是一个基于频域的方法，能够同时处理所有特征点，并对噪声和野值有良好的鲁棒性。

随机抽样一致算法—RANSAC (Random Sample Consensus)^[62]，能够解释和消除包含较大比例错误数据的实验数据集，是建模的鲁棒匹配算法，对于实验数据中的野值 (outlier) 和错误数据有良好的相容性。由于集合 T 中存在着许多的误匹配，因而直接使用 8 点算法不会取得好的效果^[61]，而鲁棒的 RANSAC 算法是解决这个问题最合适的方法^{[3][4]}。

RANSAC 算法与通常的平滑技术相反，不是用尽可能多的点去获得一个初始解

并在以后消去无效点, 而是使用满足可行条件的尽可能少的初始数据集并在可能时用一致性数据集扩大它。

通过一个简单的例子说明, 估计一组 2 维点的直线拟合。它可以被认为是对直线上的对应点作 1 维仿射变换 ($y=ax+b$) 的估计。该算法思想非常简单: 随机选择两点, 这两点定义一条直线, 到该直线的距离在一定阈值内的点数称为该线的支集。令这样的随机选择重复多次, 具有最大支集的直线就被称为是鲁棒拟合的结果。在距离阈值内的点称为内点 (inlier) (并组成一致集), 否则, 称为野值 (outlier)。直觉上, 如果直线中某一点是野值, 那么该线将不会赢得最大的支集。其中距离阈值 t 的选取和采样次数 N (即重复选择的次数) 的确定可按下述方式进行。

a) 距离阈值 t 的选取

如果测量误差服从均值为 0, 方差为 σ 的高斯分布, 则点的距离的平方 d_{\perp}^2 是高斯变量的平方和并服从一个自由度为 m 的 χ^2_m 分布, 其中 m 等于模型的余维度。对于直线, 余维度是 1—仅测量到直线的距离。如果模型是一个点, 其余维度是 2, 距离的平方是 $\frac{1}{k^2}x^2 + y^2$ 测量误差的平方和。随机变量 χ^2_m 的值小于 k^2 的概率由累积分布 $F_m(k^2) = \int_0^{k^2} \chi^2_m(\xi) d\xi$ 给出, 由该累积分布可知

$$\begin{cases} \text{内点} & d_{\perp}^2 < t^2 \\ \text{野值} & d_{\perp}^2 \geq t^2 \end{cases} \quad \text{且 } t^2 = F_m^{-1}(a)\sigma^2 \quad (4.40)$$

通常 a 取 0.95, 即点为内点的概率为 95%, 对于求两幅图像之间的单应, 距离阈值 t 可取为 $t = 5.99\sigma$ 。

b) 采样次数 N 的确定

尝试每个可能的样本通常在计算上不可行也不必要。其实只要采样次数 N 足够大, 以保证由 s 个点组成的随机样本中至少有一次没有野值的概率达到某一给定值 p 即可, 通常 p 取为 0.99。假定 w 是任意选择的数据点为内点的概率, 那么 $\varepsilon = 1 - w$ 是其为野值的概率。那么至少需要选择的次数 N 满足 $(1 - w^s)^N = 1 - p$ (s 为每次选的点数, 即样本大小), 从而

$$N = \log(1 - p) / \log(1 - (1 - \varepsilon)^s) \quad (4.41)$$

由式 (4.41), 所需的采样次数与野值所占的比例不只与其数目相关, 它还随着样本大小的增大而增大, 并且野值所占比例越大, 最低采样次数与样本大小的比例也就越大。因此在实际计算中, 为了节省运算量, 宜采用最小样本数。

c) 自适应地决定采样次数

通常, 数据中野值所占比例 ϵ 是未知的。对此情形, 算法从最坏的估计开始, 当发现更大的一致集时就把原估计更新, 通过一致集探索数据的思想可以重复应用以便自适应地确定采样次数 N 。对于每个样本, 一旦发现其一致集的 ϵ 低于当前的估计就重复更新过程, N 就再一次减少。一旦完成 N 次采样, 算法就终止。

3) 迭代修正

最后, 初始匹配集 T 中的第三类错误可以利用已经确定的对极约束和灰度约束再次寻找图像中的对应特征点交加入 T 中^[3]。

4.5 实验及结果

4.5.1 硬件设备

拍照设备使用 Canon A610 普通数码相机, 该相机 CCD 为 500 万像素, 4 倍光学变焦镜头, 镜头参数为: 7.3-29.2mm 1:2.8-4.1。

运行软件的计算机为华硕 A6000 笔记本电脑, CPU 为 Intel 移动赛扬 1.4GHz, 内存 512M, 其中集成显卡共享 8M 内存, 主机实际可用内存为 504M。

4.5.2 软件环境

实验计算机安装 Windows XP 专业版本操作系统, 版本 2002, SP2。使用 Photoshop Ver7.1 预处理图像。基于计算机视觉的三维测量实验程序使用 CVSuite Version 1.5。

CVSuite 程序由香港中文大学 ChenLEI 编写, 此软件基于计算机视觉三维测量理论, 能够通过对同一场景在不同视角下拍摄的三幅图像, 完成对真实场景的三维重建。其三维测量处理过程和相关算法如表 4-1 所述。

表 4-1 CVSuite 软件功能与算法

步骤	处理	算法	说明
1	导入图像	图像文件打开、显示	导入同一场景不同视角的三幅图像
2	提取特征点	Harris 特征点检测	提取每幅图像中的特征点
3	图像匹配	Zhang 的基于极线约束 匹配方法、松弛法	将三幅图像中的特征点进行匹配, 丢弃未匹配的特征点

4	去除野值	随机抽样一致算法 (RANSAC)、捆绑调整	将上一步中匹配的特征点集进一步求精, 得到最终的匹配点集
5	摄像机自标定	基于 Kruppa 方程的摄像机自标定	最终的匹配点集进行摄像机的自标定, 解出摄像机投影矩阵参数
6	三维重建与显示	特征点三维坐标计算、点云三角化、OpenGL 三维显示、OpenGL 纹理贴图	计算每个特征点的三维坐标、使用 OpenGL 进行三维显示

4.5.3 实验过程与数据

以下按处理顺序给出实验过程和相关的实验数据。

1) 图像获取与预处理

实验使用 Canon A610 数码相机以 300 万像素大小进行拍摄, 拍摄目标是手掌的三幅图像。三幅图像拍摄的过程中摄像机固定不动, 焦距、光圈等参数均不变 (需要关闭相机内置软件中的自动对焦功能), 将拍摄的手掌放在摄像机镜头前, 以中指为旋转轴, 以较小的角度 (使图像的视差小, 便于图像匹配处理, 角度值小于 18°) 向外旋转手掌进行拍摄。

为便于实验处理 (降低计算量, 节省实验时间), 在 Photoshop 中将图像大小缩小到 320×240 , 并将背景全部处理为蓝色, 只保留手掌部分 (消除背景在图像匹配时造成不必要的计算负担)。得到三幅不同角度的手掌图像如图 4-5 中的 (a)、(b)、(c)。

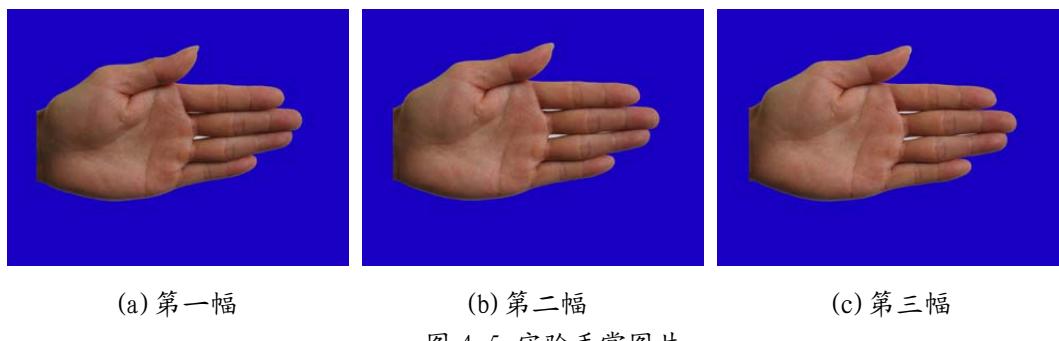


图 4-5 实验手掌图片

2) 特征点提取

将手掌图片导入后使用 Harris 算子进行特征点提取, 当对常系数 k 取不同值进行实验时, 特征点的提取数量 Cn 和分布结果并不相同, 以图像(a)为例说明特征点提取结果, 如下图所示。

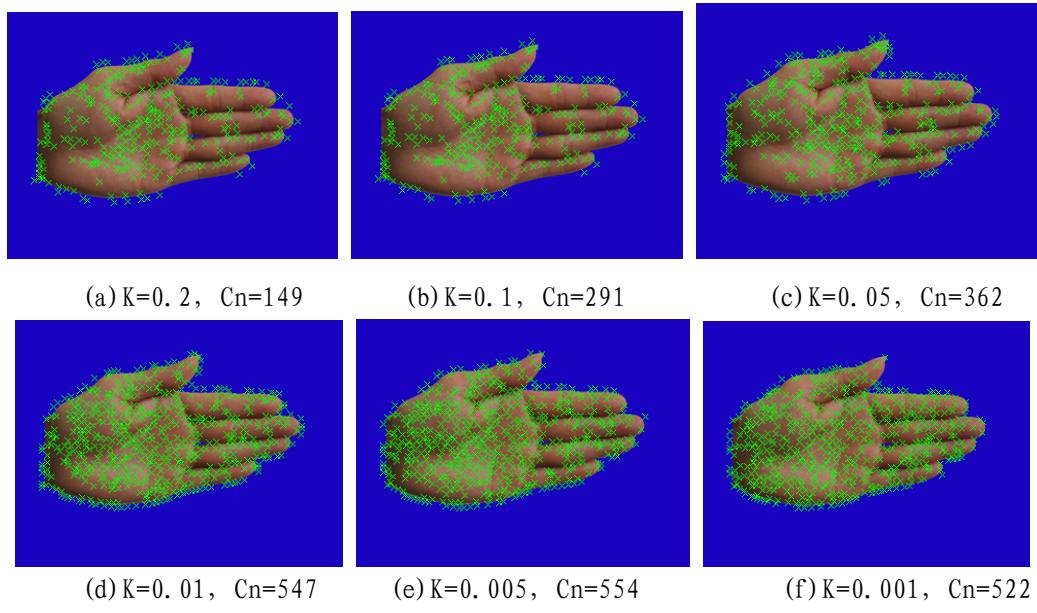


图 4-6 对第一幅图像特征点的提取

k 值与特征点数目的关系如图 4-7 所示。当 k 的取值为 0.005 时特征点数目最多, 且特征点的分布也比较合理。更多的特征点能够在三维重建中获得更精确的结果, 因此在后续实验中均以 $k=0.005$ 为特征点提取参数。

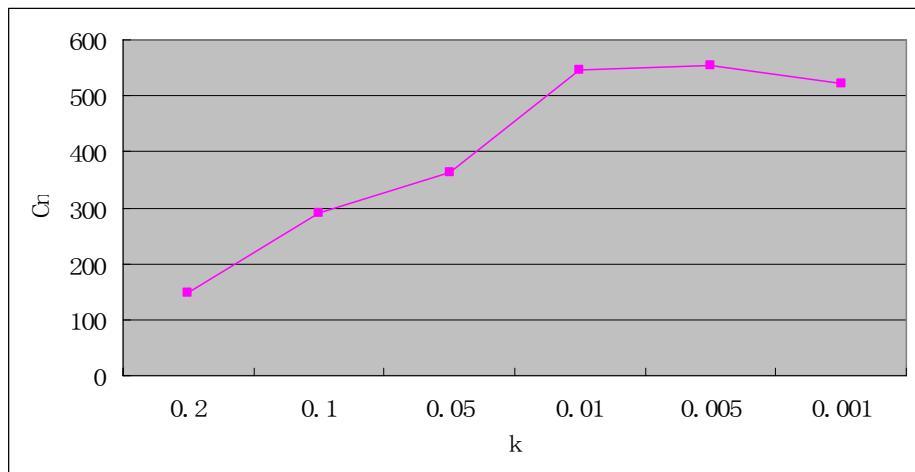


图 4-7 Harris 方法 k 值的选取与特征点数量关系

以 $k=0.005$ 为参数提取三幅图像的特征点结果如下图

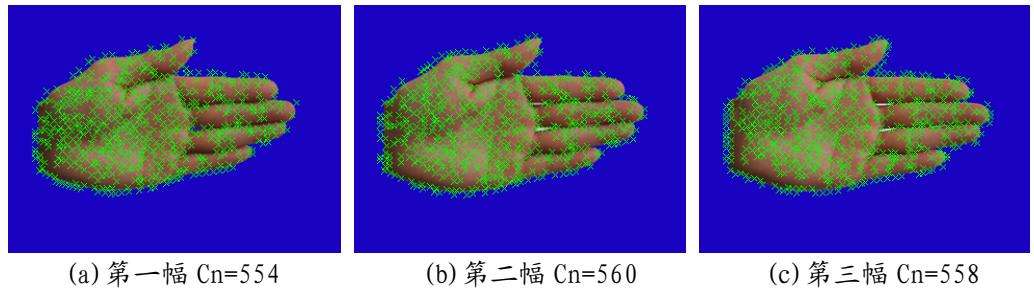
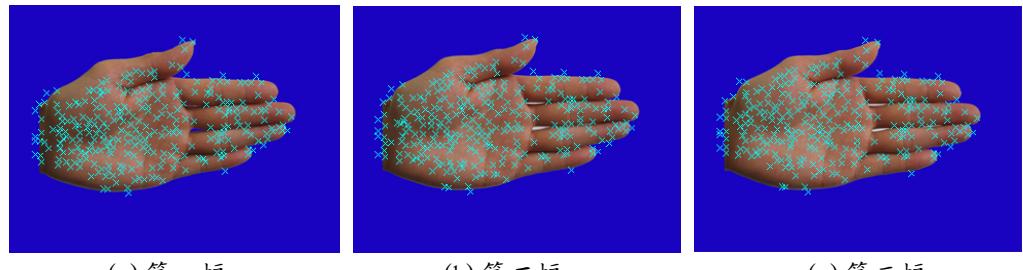


图 4-8 特征点提取结果

3) 图像匹配

Zhang 的方法参数为：窗口大小=7，阈值=0.68，搜索尺度=0.25。

松弛算法参数为：匹配强度阈值=1.25，接受率=0.75

图 4-9 图像匹配结果 (匹配点数目 $F_p=272$)

其中，(a)、(b) 两幅图像有 365 对对应特征点，(b)、(c) 两幅图像也是 365 对特征点，最终匹配点数目为 272 个，匹配的特征点主要集中在手掌中部。

4) 去除野值

去除野值后的结果如下图

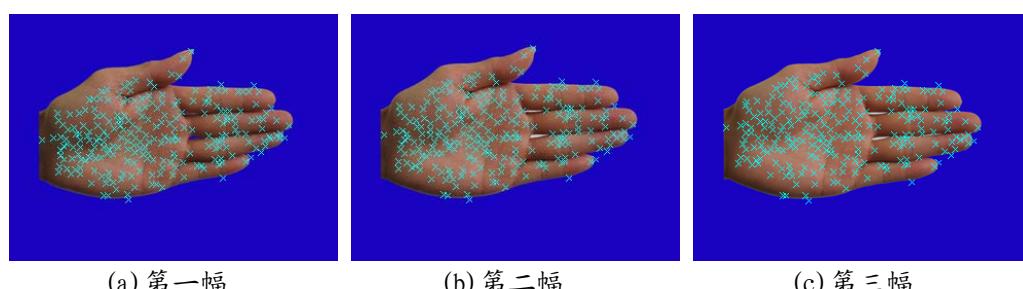


图 4-10 去除野值后内点

经过随机抽样一致算法计算出基本矩阵并通过极线约束去除掉 36 个野值点 (outliers)，剩余 236 个内点 (inliers)。

5) 摄像机自标定

采用基于 GA-Kruppa 方程的摄像机自标定结果为：

$$f_u = 151.588531, \quad f_v = 151.574350,$$

$$u_0 = -38.347518, \quad v_0 = -17.309122,$$

$$S = 24.944395$$

6) 三维重建与显示

完成以上步骤后即可计算出所有特征点的三维坐标并显示，以下为实验图片。

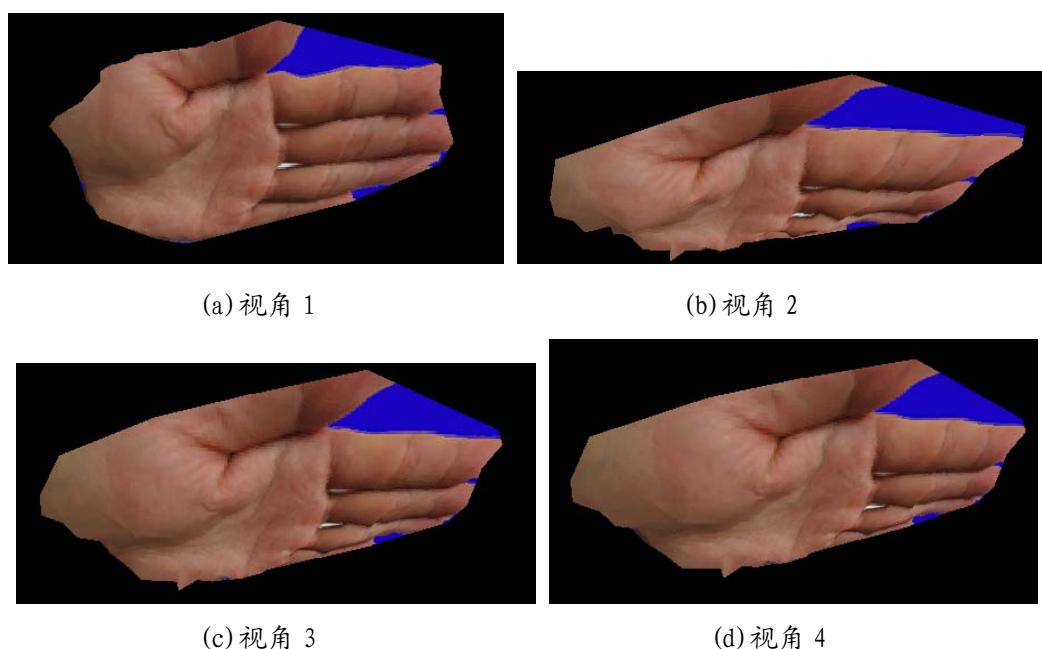


图 4-11 三维重建后贴纹理的 3D 显示

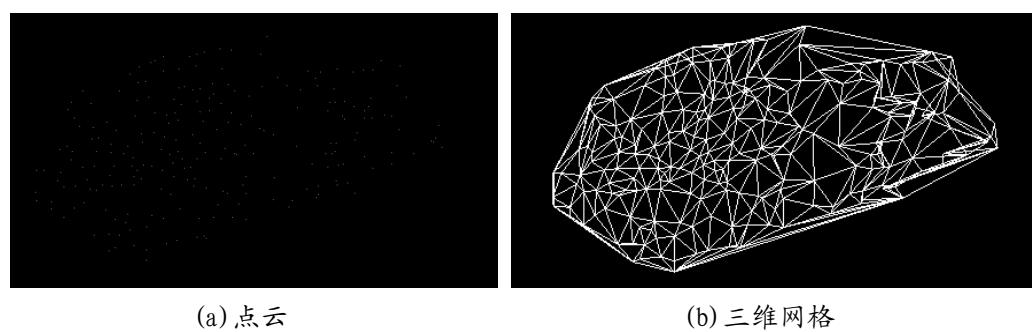


图 4-12 三维点云和三维网格

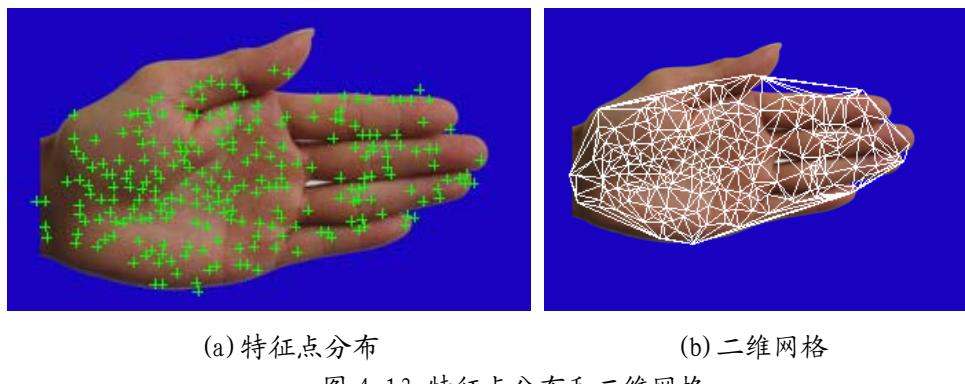


图 4-13 特征点分布和二维网格

由图 4-11 可以看出, 重建后手掌部位的曲面轮廓基本正确, 手指部位效果不佳。

4.5.4 实验分析总结

以上实验是计算机视觉对体表的三维测量实例, 最终的实验结果并不精确。对实验进行分析, 得出结论:

特征点数目越多, 重建的模型越精确; 特征点分布越密集的部位, 重建效果越细致, 如掌心位置重建后比较细致, 而在手指和手掌边缘部位, 则重建效果不佳。

程序处理中各步骤算法的参数组合也是影响结果的一个因素, 不同参数组合适应不同类型的测量对象, 另外算法中也存在许多可以改进的地方。

曲面复杂程度也是影响重建结果的重要因素。手掌相对于人体其它部位来说(除头脸部外)结构也算比较复杂, 若是其它曲面较简单的部位, 效果会更好一些(如手掌部位的效果与手指部位的效果对比)。

导致结果不精确的原因和改进方法如下:

- 1) 改进特征点提取、图像匹配算法, 产生尽可能多的匹配特征点, 并促使特征点空间位置分布均匀化;
- 2) 改用传统方法对摄像机进行标定, 提高标定精定, 最好达到亚像素级别;
- 3) 根据人体三维测量的特点分析算法的参数, 确定最佳参数组合;
- 4) 采用测量专用摄像机拍摄照片。

4.6 本章小结

本章首先介绍了基于计算机视觉三维测量的主要过程, 在此基础上深入研究计算机视觉三维测量中的基础性问题 — 摄像机标定; 对图像物征点提取问题进

行了研究;对三维测量中的关键问题 — 图像匹配问题进行了研究。本章的最后,通过实验对算法的性能和精度进行了验证。

第五章 总结与展望

本文从计算机视觉及测量技术的发展背景开始，研究了基于计算机视觉和三维测量技术的相关理论，并参照 Marr 的视觉系统框架的由理论到算法，再到物理实现的研究模型，较为系统地分析研究了基于计算机视觉的三维测量技术。

本论文在基于几何的计算机视觉的理论和算法方面进行了一些探讨，但还有很多问题需要进一步的研究和完善。

本论文的研究工作为完成最终研究目标：构建先进、完整的计算机视觉三维测量系统，打下了一定的基础，但相对于研究目标来说，目前的研究工作还只是基础性研究，离形成成果还有很远的距离。对于这个研究目标而言，仍有许多问题需要进一步深入研究。如快速而精确计算特征点三维坐标、特征点三维融合及三角化、表面纹理映射与实时三维显示等问题。

通过研究，让我深刻认识到，科学研究是一项艰苦而快乐的工作，在不断地艰难探索的同时，能够体会到获得知识的快乐。在将来的学习和工作中会有更多的问题需要我们去研究，通过不懈地学习研究，在科学的研究的道路上将会获得更丰硕的成果。

参考文献

- [1] 郑南宁. 计算机视觉与模式识别, 国防工业出版社, 1998
- [2] David Marr 著. 姚国正等译. 视觉计算理论. 北京: 科学出版社, 1988
- [3] 王文格. 基于计算机视觉的大型复杂曲面三维测量关键技术研究. [博士学位论文]. 湖南大学, 2003
- [4] 章权兵. 基于多视图的三维重构研究. [博士学位论文]. 安徽大学, 2004
- [5] Richard Hartley, Andrew Zisserman 著. 韦穗, 杨尚骏等译. 计算机视觉中的多视图几何. 安徽大学出版社, 2002. 8
- [6] 张经纬, 张青等. 人体体表损伤面积及关节活动角度智能测量与辅助鉴定系统开发与应用. 中国法医学杂志, 1999, 第 14 卷第 2 期
- [7] 梅丽, 鲍虎军, 郑文庭等. 基于实拍图像的人脸真实感重建. 计算机学报, 2000, 第 23 卷第 9 期
- [8] 李健, 刘先林, 刘凤德. 基于双立体同步摄影的面部三维重建与量测. 测绘科学, 2002, 第 27 卷第 2 期
- [9] 吴剑波, 骆清铭, 谭玉山. 网格投影式立体视觉三维表面重建系统. 光电工程, 2004, 第 31 卷第 11 期
- [10] 徐常胜, 徐众, 田英利等. 三维面像数据采集和重建系统. 中国图象图形学报, 1998, 第 3 卷第 2 期
- [11] T. Varady, R. R. Martin, J. Cox. Reverse engineering of geometric models – an introduction. Computer Aided Design, 1997, 29 (4):25 5–268
- [12] Q. J. Peng, M. Loftus. A new approach to reverse engineering based on vision information. International Journal of Machine Tools & Manufacture, 1998, 38:881–899
- [13] 于起峰, 陆宏伟, 刘肖琳. 基于图像的精密测量与运动测量. 北京: 科学出版社, 2002
- [14] 王之卓. 摄影测量原理续编, 测绘出版社, 1986
- [15] F. Kobayashi, T. Fukuda, K. Shimojima, T. Takusagawa. Shape measurement method integrating stereo vision and shape-from-shading with evolutionary programming. In:Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems, 1997
- [16] S. Okada, M. Imade, H. Miyauchi. 3-D shape measurement of free-form machined surfaces by optical ring imaging system. In:Proceedings of

- the 24th Annual Conference of the IEEE. 1998
- [17] 徐常胜, 徐众, 田英利等. 三维面像数据采集和重建系统. 中国图象图形学报, 1998, 第 3 卷第 2 期
- [18] T. N. Tan, Efficient Image Gradient Based Vehicle Localisation. IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 9, NO. 8, AUGUST 2000
- [19] 朱若磊. 基于计算机视觉的高速焊缝跟踪系统的研究. [博士学位论文]. 华南理工大学, 2004
- [20] 陈莹, 潘俊民. 大型物料堆体积的计算机视觉测量方法. 上海交通大学学报, 2002, 第 7 期
- [21] 孙怡. 人体运动跟踪的方法与实验研究. [博士学位论文]. 大连理工大学, 2002
- [22] 张春森. 序列立体图象三维运动物体定位与跟踪. [博士学位论文]. 武汉大学, 2004
- [23] 皮文凯. 基于全方位视觉的人体运动检测与跟踪. [硕士学位论文]. 北京大学, 2004
- [24] 冯文灏. 近景摄影测量的基本技术提要. 测绘科学, 2000, 第 25 卷第 4 期
- [25] 程效军, 罗武. 基于非量测数字相机的近景摄影测量. 铁路航测, 2002 第 28 卷第 1 期
- [26] 邓钢. 非量测影像数字化近景摄影测量的部分问题探讨. 昆明理工大学学报, 2002, 第 27 卷 第 04 期
- [27] 贺跃光, 王秀美, 曾卓乔. 数字化近景摄影测量系统及其应用. 矿冶工程, 2001, 第 4 期
- [28] 周富强. 双目立体视觉检测的关键技术研究. [博士后士]. 北京航空航天大学, 2002
- [29] 王冬林. 车身覆盖件曲面延展设计方法的实现及应用研究. [硕士学位论文]. 吉林大学, 2001
- [30] K. I. Boyer, A . CK AK. Color-encoded structured light for rapid active ranging. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1987, Volum 9, No. 1
- [31] T. Pancewicz, M. Kujawinska. CAD/CAM/CAE representation of 3D objects measured by fringe projection, In: Proc. SPIE 3479, 1998
- [32] G. Sansoni, M. Carocci, R. Rodella. Calibration and performance evaluation of a 3-D imaging sensor based on the projection of structured light, IEEE Trans. Instr. Meas. 2000

- [33] Rogen Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE JOURNAL OF ROBOTICS AND AUTOMATION*, VOL. RA-3, NO. 4, AUGUST 1987
- [34] J. Y. Weng, P. Cohen, M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1992, 14
- [35] Kin Hong Wong, Michael Ming Yuen Chang. 3D Model Reconstruction by Constrained Bundle Adjustment. *ICPR 2004*, Cambridge, Oral presentation paper
- [36] Berthold K. P. Horn. Tsai's camera calibration method revisited. *MIT Open Course Ware*. 2000
- [37] Motilal Agrawal, Larry S. Davis. Complete camera calibration using spheres: A dual-space approach. *Department of Computer Science, University of Maryland*. 2002
- [38] Fay Huang, Shou-Kang Wei, Reinhard Klette. Comparative Studies of Line-based Panoramic Camera Calibration. *Inst. of Information Science Academia Sinica Taipei, Taiwan, Dep. of Information and Computer Science Keio University Yokohama, Japan, CITR, Computer Science Dep. The University of Auckland Auckland, New Zealand*. 2003
- [39] H. Zollner, R. Sablatnig. Comparison of Methods for Geometric Camera Calibration using Planar Calibration Target. *Pattern and Image Processing Group (PRIP) Vienna University of Technology*
- [40] Zhengyou Zhang. A Flexible New Technique for Camera Calibration. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, VOL. 22, NO. 11, NOVEMBER 2000
- [41] Chang Shu, Alan Brunton, Mark Fiala. CAMcal: A Program for Camera Calibration Using Checkerboard Patterns. *Institute for Information Technology National Research Council Canada*. 2001
- [42] Xiaoqiao Meng, Hua Li and Zhanyi Hu. A New Easy Camera Calibration Technique Based on Circular Points. *National Laboratory of Pattern Recognition (NLPR) Chinese Academy of Sciences (CAS), Journal of Software*. Vol. 13, No. 5, 2002
- [43] Carlo Colombo, Dario Comanducci, and Alberto Del Bimbo. Camera

- calibration with two arbitrary coaxial circles. Dipartimento di Sistemi e Informatica Via S. Marta 3, 50139 Firenze, Italy. 2006
- [44] Francisco Abad, Emilio Camahort, and Roberto Viv. Camera Calibration Using Two Concentric Circles. Universidad Polit'ecnica de Valencia. ICIAR, 2004
- [45] Qurban Memon, Sohaib Khan. Camera calibration and three-dimensional world reconstruction of stereo-vision using neural networks. International Journal of Systems Science, 2001, volume 32, number 9
- [46] Xianghua Ying, Zhanyi Hu. Catadioptric Camera Calibration Using Geometric Invariants. In Proc. of International Conference on Computer Vision, 2003
- [47] 邱茂林, 马颂德, 李毅. 计算机视觉中的摄像机定标综述. 自动化学报, 2000, 26(1)
- [48] 吴福朝, 阮宗才, 胡占义. 非线性模型下的摄像机自标定. 计算机学报, 2002, 25(3)
- [49] Fabio Remondino, Niclas B'orlin. PHOTOGRAHMETRIC CALIBRATION OF IMAGE SEQUENCES ACQUIRED WITH A ROTATING CAMERA. Institute for Geodesy and Photogrammetry, ETH Zurich, Switzerland, Department of Computing Science, Umea University, Sweden
- [50] 陈泽志, 吴成柯. 一种新的手提相机自定标方法. 中国图象图形学报, 2003, 第3期
- [51] Z. Zhang, R. Deriche, O. Faugeras et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence, 1995, 75
- [52] 徐建华. 图像处理与分析. 北京: 科学出版社, 1992
- [53] 谢东海, 詹总谦, 江万寿. 改进 Harris 算子用于点特征的精确定位. 测绘信息与工程, 2003, 28(2)
- [54] 管海燕, 郭建星. 常用图像边缘检测算子定位精度对比研究. 测绘与空间地理信息, 2005
- [55] 何凯, 安如, 周绍光, 金夏玲. 一种快速角点探测算子研究. 测绘学报, 2005(3)
- [56] P. Pritchett, A. Zisserman. Wide baseline stereo matching. In: Proc. ICCV, 1998

- [57] M. Lhuillier, L. Quan. Robust dense matching using local and global geometric constraints. In: Proc. ICPR, Barcelona, 2000
- [58] C. Schmid, R. Mohr. Local greyvalue invariants for image retrieval. PAMI, 1997
- [59] T. Tuytelaars, L. van Gool. Matching of affinely invariant regions for visual servoing. In: Proc. ICRA, 1999
- [60] D. Tell, S. Carlsson. Wide baseline point matching using affine invariants computed from intensity profiles. In : Proc. ECCV, 2000
- [61] P. Torr, D .M urray. The development and comparison of robust methods for estimating the fundamental matrix. Internationa lournal of Computer Vision, 1997
- [62] M. A. Fischler, R. C. B olles. Random sample consensus: Aparadigm for model fitting with applications to image analysis and automated cartography. GIP, 1981
- [63] Zhengyou Zhang, Charles Loop. Estimating the Fundamental Matrix by Transforming Image Points in Projective Space. Computer Vision and Image Understanding, 2001, Vol. 82, No. 2
- [64] Stefan Lehmann, Vaughan L. Clarkson, Andrew P. Bradley et al. Robust Fundamental Matrix Determination without Correspondences. Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP)

致 谢

感谢我的导师何继善老师在研究生阶段对我的的关怀和指导。感谢熊平老师在我学习研究时期给我的指导和帮助，让我了解到法医学领域存在的研究课题，并最终确定论文题目。

感谢在信息物理工程学院生物医学工程实验室（地学楼226室）工作过的老师，和一起研究学习过的同学，在我们共同学习和研究的过程中，加深了对生物医学研究领域的认识，让我接触和学习到生物医学相关的各方面知识，开阔了我的视野和专业知识面。在226实验室中，大家营造了团结、向上、勤奋、务实的和谐气氛，使我在一个轻松而紧张的环境中学习和研究。

赵于前老师、徐大宏博士在学术研究方面曾给予我大力的支持，在此表示感谢。傅国红、徐嘉瑞、肖嘉莹、罗东礼、邓素平、王飞等同学对我在研究工作和生活方面的帮助也让我的研究工作开展地更加顺利，在此表示衷心的感谢。

最后感谢我的家人、女友和朋友，是他们始终在背后的默默支持和奉献，使得我在遇到困难的时候，始终有继续坚持的勇气和信心。

作者攻读硕士学位期间的发表论文情况

发表的论文

1. 张必明, 熊平, 肖嘉莹. 多边形分解方法计算人体体表局部损伤面积, 中国法医学杂志, 2006. 4, 已录用。

参与的研究项目

1. 第四军医大可视人切片三维重建项目
2. 法医临床鉴定体表损伤测量项目